

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2004, 14:42:06 ; Search time 79 Seconds

(without alignments)
2929.409 Million cell updates/sec

Title: US-10-054-691-2

Perfect score: 7766
Sequence: 1 MGRPGIFILELLILGQGT.....RCRRGRREDPPMSRLVAL 1458

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

A_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	7766	100.0	1458	23	ABB09555	Human lipase NHL
2	7762	99.9	1458	23	ABB09556	Human lipase NHL
3	7759	99.9	1458	24	AAE34440	Human lipid-associ
4	7383.5	95.1	1433	23	ABU65083	Human NOV24c prote
5	6311.5	81.3	1419	23	ABU65081	Human NOV24a prote
6	6253.5	80.5	1216	23	AAE22860	Human phospholipas
7	5279.5	68.0	1450	18	AAW30751	Human phospholipas
8	4911	63.2	969	24	AAE34448	Rat phospholipase-
9	2492	32.1	472	23	ABP53556	Human lipid-associ
						Human phospholipas

10	1525	19.6	310	23	ABU65082 Human NOV24b prote
11	1360	17.5	267	22	ABR11053 Human phospholipas
12	1360	17.5	267	22	AAW25824 Human protein sequ
13	584	7.5	424	22	ABR71556 Drosophila melanog
14	560	7.2	148	22	ABR96111 Human testicular a
15	560	7.2	148	22	AAW95420 Human reproductive
16	527	6.8	101	23	ABR89352 Human polypeptide
17	515	6.6	447	22	ABR65406 Drosophila melanog
18	459	5.9	132	22	ABR11237 Human phospholipas
19	324.5	4.2	334	22	ABG02997 Novel human diagno
20	249	3.2	109	22	AAE10214 Human bone marrow
21	178.5	2.3	102	23	ABP34826 Human ORF3799 prot
22	138	1.8	10182	23	ABP38314 Staphylococcus epi
23	137.5	1.8	639	22	ABR68805 Drosophila melanog
24	137.5	1.8	1369	22	ABR58070 Drosophila melanog
25	136	1.8	1424	23	ABG66726 Human novel polype
26	136	1.8	1455	19	AAW48663 Fancconi anaemia of
27	132.5	1.7	2542	22	ABR71137 Drosophila melanog
28	129.5	1.7	1218	22	AAE62801 DNA ligase II amin
29	128	1.6	589	23	ABR53511 Lactococcus lactis
30	127.5	1.6	1016	23	AAW41312 CF-5 pathogen resi
31	127	1.6	787	24	ABJ26529 Aspergillus fumiga
32	125.5	1.6	1339	23	ABR06077 Human NS protein s
33	125	1.6	30	18	AAW30753 Phospholipase-B/11
34	124	1.6	1178	18	AAW30763 Mannose-1-phosphat
35	124	1.6	2828	23	ABG32897 Human Adlican prot
36	124	1.6	2828	23	ABG31319 Human Adlican prot
37	124	1.6	2828	23	ABG31352 Human Adlican-2 pr
38	124	1.6	2828	24	ABR7459 Human breast cancer
39	124	1.6	2828	24	ABJ37051 Human breast cancer
40	124	1.6	2828	24	ABU56609 Lung cancer-associ
41	123.5	1.6	1230	19	AAW98275 H. pylori GHP0 690
42	122	1.6	744	22	ABR71226 Drosophila melanog
43	121.5	1.6	1031	22	ABR65127 Drosophila melanog
44	121.5	1.6	1097	22	ABR61579 Drosophila melanog
45	121.5	1.6	2597	21	AAV53664 Mechanical stress

ALIGNMENTS

RESULT 1	ABB09555 standard; Protein; 1458 AA.
ID	ABB09555
XX	ABB09555;
AC	21-OCT-2002 (first entry)
XX	
DE	Human Lipase NHL (A1a 1318 variant).
XX	
XX	Human, lipase; chromosome 2; phospholipase B homologue; thyroid; brain;
KW	inflammatory disease; proliferative disease; infectious disease;
KW	clotting disorder; cancer; drug screening; mental disorder; NHL;
KW	diagnostic reagent; clinical trial monitoring; cosmetic; nutritional;
KW	mutation detection; gene expression analysis; transgenic animal;
KW	neurotic; cytostatic; antiinflammatory; single nucleotide polymorphism;
KW	SNP; gene therapy; enzyme.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	Misc-difference 1318
FT	Location/Qualifiers
FT	/note="Val replaces Ala in a polymorphic variant (ABB09556)"
PN	W0200259328-A1.
XX	
PD	01-AUG-2002.
XX	
PF	22-JAN-2002; 2002MO-US01715.
XX	
PR	24-JAN-2001; 2001US-264049P.

XX (LEXI-) LEXICON GENETICS INC.
PA Yu X, Miranda M, Turner CA;
PI WPI; 2002-599797/64.
XX N-PSDB; ABQ77623.
PT Polynucleotides encoding human lipases that are structurally related to
PT animal lipases, particularly phospholipase B, useful for drug
PT screening, diagnosis and in gene therapy of biological disorders -
XX Claim 2; Page 37-41; 44pp; English.
PS
XX The invention relates to a novel human lipase (NHL; ABB09555, ABB09556)
CC and to nucleic acids encoding it (ABQ77623, ABQ77624). The NHL has
CC structural similarity with animal lipases, particularly phospholipase B.
CC Polynucleotides encoding NHL were obtained using human genomic sequences
CC in conjunction with human thyroid and brain cDNAs. The NHL gene is
CC located on chromosome 2, and contains a C/T polymorphism at position
CC 3953 of the open reading frame (ORF), resulting in an Ala/Val
CC substitution at position 1318 in the protein. NHL nucleotides and
CC proteins are useful for treating disorders such as inflammatory or
CC proliferative disease, infectious disease, clotting disorders, and
CC cancer. They can also be used in screening for compounds useful in
CC the treatment of mental, biological or medical disorders, as diagnostic
CC reagents, in clinical trial monitoring and in cosmetic and nutritional
CC applications. NHL nucleotides can additionally be used in the detection
CC of disease-associated mutations, in the analysis of gene expression, for
CC the recombinant expression of NHL, to generate transgenic animals, in
CC gene therapy, and as part of ribozyme and/or triple helix sequences
CC useful in the modulation of NHL gene expression. The present sequence
CC represents the Ala 1318 variant of NHL.

SQ Sequence 1458 AA;

Query Match 100.0%; Score 7766; DB 23; Length 1458;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;

Matches 1458; Conservative 0; Indels 0; Gaps 0;

QY 1 MGLRPGIFLLELLLLGQGTPOIHTSPKSTLSEGLWPELTKNSPPPCNPNKLGVMNSK 60
DB 1 MGLRPGIFLLELLLLGQGTPOIHTSPKSTLSEGLWPELTKNSPPPCNPNKLGVMNSK 60
QY 61 SVHSLSKPSDIFKFAAIGNLEIPDPGTCDELEKQDWTERRPOQVCMGVNTVSDIIRYSPS 120
DB 61 SVHSLSKPSDIFKFAAIGNLEIPDPGTCDELEKQDWTERRPOQVCMGVNTVSDIIRYSPS 120
QY 121 VPMPCVCHTGRVIPHDAEDLMTIOAELVNMKENLQDLPQPMKLNVEFSNASQCYLC 180
DB 121 VPMPCVCHTGRVIPHDAEDLMTIOAELVNMKENLQDLPQPMKLNVEFSNASQCYLC 180
QY 181 PSAQNGLAAGVDELWGLVDYLQOEVPRAFVNLVDLSEVAEVSROYHGTWLSPAPEPCN 240
DB 181 PSAQNGLAAGVDELWGLVDYLQOEVPRAFVNLVDLSEVAEVSROYHGTWLSPAPEPCN 240
QY 241 CSEETTRILAVVQMSQZAMNSILASRSYSEQSFVVVQPFYETTPSLHSEDPLOD 300
DB 241 CSEETTRILAVVQMSQZAMNSILASRSYSEQSFVVVQPFYETTPSLHSEDPLOD 300
QY 301 STTIAHLNMRMMEPAGEKDEPLSVKHGRPMKCPQSPLYLFSYRNSNYLTRLOKPDKL 360
DB 301 STTIAHLNMRMMEPAGEKDEPLSVKHGRPMKCPQSPLYLFSYRNSNYLTRLOKPDKL 360
QY 361 EVREGAIRCPDPSDPTVPTSVHRLKPADINVIGALGDSLTAGNGAGSTPGANVLDVLTQ 420
DB 361 EVREGAIRCPDPSDPTVPTSVHRLKPADINVIGALGDSLTAGNGAGSTPGANVLDVLTQ 420
QY 421 YRGLSNVSGDENIGTITTLANTLREFNPSLKGFSVGTGKETSBNALNOAVAGRAEDL 480
DB 421 YRGLSNVSGDENIGTITTLANTLREFNPSLKGFSVGTGKETSBNALNOAVAGRAEDL 480
QY 481 PVQARLVLDLMDKNDTRHFOEDMKITLFLIGNDLDCDFCDVLYHYSQNTDNI GKALDI 540

DB 481 PVQARLVLDLMDKNDTRHFOEDMKITLFLIGNDLDCDFCDVLYHYSQNTDNI GKALDI 540
QY 541 LHAEPRAFVNLVTVLEIVNLREL YOEKKVYCPRMILRSICPCVLYKFDNSTELATLIEF 600
DB 541 LHAEPRAFVNLVTVLEIVNLREL YOEKKVYCPRMILRSICPCVLYKFDNSTELATLIEF 600
QY 601 NKKFOEKTQILISGSDYDREDFTVVQPFENVMPKSEGLPDNSFFAPDCFFSSKS 660
DB 601 NKKFOEKTQILISGSDYDREDFTVVQPFENVMPKSEGLPDNSFFAPDCFFSSKS 660
QY 661 HSRBAASLMMNNMLEPVQOKTTRKFKENKINITCPNOVQPLRTYKMSQGHGTWLPCKDR 720
DB 661 HSRBAASLMMNNMLEPVQOKTTRKFKENKINITCPNOVQPLRTYKMSQGHGTWLPCKDR 720
QY 721 APSALHPTSVAHARPADIQVVAALGSLTRNGNGISKPPDLPVTTQYRGLSAGSDGS 780
DB 721 APSALHPTSVAHARPADIQVVAALGSLTRNGNGISKPPDLPVTTQYRGLSAGSDGS 780
QY 781 LEVNTTLPLNLRBFNRLTGYAVGTGDANDTNAPLNOAVGAKAEDLMSOVOTLMQMKD 840
DB 781 LEVNTTLPLNLRBFNRLTGYAVGTGDANDTNAPLNOAVGAKAEDLMSOVOTLMQMKD 840
QY 841 DHRVNFHEDKVITVVLIGSDLCDYCTDSNLYSAANFVDHLARNALDVLHREVPRLVNLV 900
DB 841 DHRVNFHEDKVITVVLIGSDLCDYCTDSNLYSAANFVDHLARNALDVLHREVPRLVNLV 900
QY 901 DFLNPTIMROVFLGNPDKCPVQOASVLCNCVLTLENSOELARLEAFSRAYSKRELVG 960
DB 901 DFLNPTIMROVFLGNPDKCPVQOASVLCNCVLTLENSOELARLEAFSRAYSKRELVG 960
QY 961 SGRVYDQEDSVYLQPFQNTICLPVLDLGLPDSFPAPDCIHNOKFHSGLAALMTNML 1020
DB 961 SGRVYDQEDSVYLQPFQNTICLPVLDLGLPDSFPAPDCIHNOKFHSGLAALMTNML 1020
QY 1021 EPLGSKTETLDLRAENPITCPTONEPFLRTPRNSNTYPIKPAIENWGSDFLCTBWKASN 1080
DB 1021 EPLGSKTETLDLRAENPITCPTONEPFLRTPRNSNTYPIKPAIENWGSDFLCTBWKASN 1080
QY 1081 SVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSDDLPTSMRGLSWSIGDGNLETHYT 1140
DB 1081 SVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSDDLPTSMRGLSWSIGDGNLETHYT 1140
QY 1141 LPMILKKFNPVYLIGFSTWEGTAGLVNVAEGARADMPOANDLVERMKNSPDINLEKD 1200
DB 1141 LPMILKKFNPVYLIGFSTWEGTAGLVNVAEGARADMPOANDLVERMKNSPDINLEKD 1200
QY 1201 MKLVTLFLIGVNDLCHYCENPEALATEYVOHIQOALDILSEELPRAFVNVEVMEIASLY 1260
DB 1201 MKLVTLFLIGVNDLCHYCENPEALATEYVOHIQOALDILSEELPRAFVNVEVMEIASLY 1260
QY 1261 OGQGGKCAMLAAONNCTCLRHSQSLEKOLKVVNNLQHGSSFSYMHQYTORDEFAVY 1320
DB 1261 OGQGGKCAMLAAONNCTCLRHSQSLEKOLKVVNNLQHGSSFSYMHQYTORDEFAVY 1320
QY 1321 VQPFQNTLPLNERDPTDITFSEDCFFHPSDGHAEALAMNNMLEPGRKTTSNFT 1380
DB 1321 VQPFQNTLPLNERDPTDITFSEDCFFHPSDGHAEALAMNNMLEPGRKTTSNFT 1380
QY 1381 HSRKALKCPSPESPYLTLRNSRLPDQAEAPVLYMVAAGVGLVGIIGTVVMRC 1440
DB 1381 HSRKALKCPSPESPYLTLRNSRLPDQAEAPVLYMVAAGVGLVGIIGTVVMRC 1440
QY 1441 RRGGRREDPMSLRTVAL 1458
DB 1441 RRGGRREDPMSLRTVAL 1458
RESULT 2
ABB09556
ID ABB09556 standard; Protein; 1458 AA.
XX
AC ABB09556;

Query Match	Beat Local Similarity	Matches 1457	Score 7759	DB 24	Length 1458
99.9%	99.9%	0	0	1	0
Conservative	0	Mismatches	1	Indels	Gaps

1 MGLARGLFLLELLLLGGCTPQIHTSPKSLTLEGLMWEETLKNSEFPNCNPKLGVNMSK 60

Db 1 MGRLPGIFLLELLLLLLGQTPQIHTSPKSTLEGLQWPELTKNSPPECNPNKLGVNPSK 60
 QY 61 SVHSLSKSDIKFYVAIGNLEIPDPGTGDELEKODMTERPOQVCGWVTSLDIIRYSPS 120
 Db 61 SVHSLSKSDIKFYVAIGNLEIPDPGTGDELEKODMTERPOQVCGWVTSLDIIRYSPS 120
 QY 121 VPMVCHTGRV1PHDGAEDLM1QAQELVNMKNELQDFOPMKLIINVEFSNASCYLC 180
 Db 121 VPMVCHTGRV1PHDGAEDLM1QAQELVNMKNELQDFOPMKLIINVEFSNASCYLC 180
 QY 181 PSQAQNGLAAGVDELMLGVLDYLOQEVPRAFVNLVDLSVAVESRQYHGHWLSPAPPCN 240
 Db 181 PSQAQNGLAAGVDELMLGVLDYLOQEVPRAFVNLVDLSVAVESRQYHGHWLSPAPPCN 240
 QY 241 CSSEETRLAYVMQWSYOZAMNSLASSRSEOSFTVVQPFPEYETPSLHSDPELOD 300
 Db 241 CSSEETRLAYVMQWSYOZAMNSLASSRSEOSFTVVQPFPEYETPSLHSDPELOD 300
 QY 301 STTLAHLNMRMMEPAGEKDEPLSVKGRPMKCPQSPYLFYSRNSNYLTRLOKPODKL 360
 Db 301 STTLAHLNMRMMEPAGEKDEPLSVKGRPMKCPQSPYLFYSRNSNYLTRLOKPODKL 360
 QY 361 EYREGAEIRCPDQPSPTVPTSVHRLKPADINVIGALGDSLTAAGNGAGSTPGVNLVLTQ 420
 Db 361 EYREGAEIRCPDQPSPTVPTSVHRLKPADINVIGALGDSLTAAGNGAGSTPGVNLVLTQ 420
 QY 421 YRGLSMSGDENIGVTTLANILREFNPSLKGSVGTGKETSNAFLNQAAGGRDEL 480
 Db 421 YRGLSMSGDENIGVTTLANILREFNPSLKGSVGTGKETSNAFLNQAAGGRDEL 480
 QY 481 PVQARRLVDMKNDTRIHFOEDWKIITLFTIGNDLDCFCNDLVHSPQNFNDINIGKALDI 540
 Db 481 PVQARRLVDMKNDTRIHFOEDWKIITLFTIGNDLDCFCNDLVHSPQNFNDINIGKALDI 540
 QY 541 LHAEPVPAFVNLVTVLEIVNLRELYQEKVYCPRMILRSICPCVLYKFDNDSTELATLIEF 600
 Db 541 LHAEPVPAFVNLVTVLEIVNLRELYQEKVYCPRMILRSICPCVLYKFDNDSTELATLIEF 600
 QY 601 NKRFQEKTHOLISGRYDTEDEFTVVVQPFPEVNDMKTSEGLPDNSFFAPDCPHFSKS 660
 Db 601 NKRFQEKTHOLISGRYDTEDEFTVVVQPFPEVNDMKTSEGLPDNSFFAPDCPHFSKS 660
 QY 661 HSRBAASALMMNMLEPVQOKTTRHKFNKINITCPNOVQPLRTYKNSMOGHGTMVLPDRDR 720
 Db 661 HSRBAASALMMNMLEPVQOKTTRHKFNKINITCPNOVQPLRTYKNSMOGHGTMVLPDRDR 720
 QY 721 APSALHPTSVHALRPADIQVVAALGDSLTAAGNGISGRDLDPVTTQYRGLSYSAGDGS 780
 Db 721 APSALHPTSVHALRPADIQVVAALGDSLTAAGNGISGRDLDPVTTQYRGLSYSAGDGS 780
 QY 781 LENVTTLPNILREFNRLTGYAVGTGDANTNAFLNQAAGCAEDLMSOVOTLMQMKD 840
 Db 781 LENVTTLPNILREFNRLTGYAVGTGDANTNAFLNQAAGCAEDLMSOVOTLMQMKD 840
 QY 841 DHRVNFHEDKVITVLLIGSDLCDYCTDSNLYSAANVVDLRYNALDYLHREVPVNLV 900
 Db 841 DHRVNFHEDKVITVLLIGSDLCDYCTDSNLYSAANVVDLRYNALDYLHREVPVNLV 900
 QY 901 DFLNPTIMRQVFLGNPDKCPVQAASVLCNCVLTILRENSQELARLEAFSRAVRSMBELV 960
 Db 901 DFLNPTIMRQVFLGNPDKCPVQAASVLCNCVLTILRENSQELARLEAFSRAVRSMBELV 960
 QY 961 SGRYDQEDSSVYLQPFQNIOLPVLADGJPDTSFFAPDCIHPQKHSQIAGALMTNML 1020
 Db 961 SGRYDQEDSSVYLQPFQNIOLPVLADGJPDTSFFAPDCIHPQKHSQIAGALMTNML 1020
 QY 1021 EPLGSKTETDLEAEMPICTPONEPFLRTPRNSNTYYP1KPA1EMNGSDFLCTEMKASN 1080
 Db 1021 EPLGSKTETDLEAEMPICTPONEPFLRTPRNSNTYYP1KPA1EMNGSDFLCTEMKASN 1080
 QY 1081 SVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSSDLP1TSMWGLSISGDGDLERTHTT 1140
 Db 1081 SVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSSDLP1TSMWGLSISGDGDLERTHTT 1140

QY 1141 LPNILLKFNBYLLGFSTSTWEGTAGLVNAEGARADMPAQANDLYERMKNSPDINLEKD 1200
 Db 1141 LPNILLKFNBYLLGFSTSTWEGTAGLVNAEGARADMPAQANDLYERMKNSPDINLEKD 1200
 QY 1201 WKLVTLFIGVNDLCHYCENBEAHLATEYVOHIOALDIISEELPRAFVNVVEWELASLY 1260
 Db 1201 WKLVTLFIGVNDLCHYCENBEAHLATEYVOHIOALDIISEELPRAFVNVVEWELASLY 1260
 QY 1261 QGGGKCAMLAANNCTCLRHSSQSLSEKOLKRYNNLQHGISFSFWHOTQOREDAVY 1320
 Db 1261 QGGGKCAMLAANNCTCLRHSSQSLSEKOLKRYNNLQHGISFSFWHOTQOREDAVY 1320
 QY 1321 VQPFQNTLTPLENGRDTLTPFSEDCPHFSDRGHAEMAIALMNNMLEPVGRKTSNNFT 1380
 Db 1321 VQPFQNTLTPLENGRDTLTPFSEDCPHFSDRGHAEMAIALMNNMLEPVGRKTSNNFT 1380
 QY 1381 HSRAKLKCPSPESPYLYTLNRSRLPDQAEAEVLYMAVPVAAGVGLVVGIICTVVMRC 1440
 Db 1381 HSRAKLKCPSPESPYLYTLNRSRLPDQAEAEVLYMAVPVAAGVGLVVGIICTVVMRC 1440
 QY 1441 RRGGRREDPMSLRITVAL 1458
 Db 1441 RRGGRREDPMSLRITVAL 1458

RESULT 4
 ABU65083
 ID ABU65083 standard; Protein; 1423 AA.
 AC ABU65083;
 XX
 DT 20-MAY-2003 (first entry)
 XX
 DE Human NOV24c protein.
 XX
 KM NOVX; cytosolic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
 KM hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 KM human.
 XX
 OS Homo sapiens.
 OS
 PN MO200272757-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 08-MAR-2002; 2002MO-US06908.
 XX
 PR 08-MAR-2001; 2001US-274101P.
 PR 08-MAR-2001; 2001US-274194P.
 PR 08-MAR-2001; 2001US-274281P.
 PR 08-MAR-2001; 2001US-274322P.
 PR 09-MAR-2001; 2001US-274849P.
 PR 12-MAR-2001; 2001US-275235P.
 PR 13-MAR-2001; 2001US-275578P.
 PR 13-MAR-2001; 2001US-275579P.
 PR 13-MAR-2001; 2001US-275601P.
 PR 14-MAR-2001; 2001US-276000P.
 PR 16-MAR-2001; 2001US-27676P.
 PR 19-MAR-2001; 2001US-276994P.
 PR 20-MAR-2001; 2001US-277239P.
 PR 20-MAR-2001; 2001US-277321P.
 PR 20-MAR-2001; 2001US-277327P.
 PR 21-MAR-2001; 2001US-277791P.
 PR 22-MAR-2001; 2001US-277833P.
 PR 23-MAR-2001; 2001US-278152P.
 PR 26-MAR-2001; 2001US-278894P.
 PR 27-MAR-2001; 2001US-278999P.
 PR 27-MAR-2001; 2001US-279036P.
 PR 28-MAR-2001; 2001US-279344P.
 PR 30-MAR-2001; 2001US-277338P.
 PR 30-MAR-2001; 2001US-279959P.
 PR 30-MAR-2001; 2001US-280233P.

PR 02-APR-2001; 2001US-280802P.
 PR 02-APR-2001; 2001US-280822P.
 PR 02-APR-2001; 2001US-280900P.
 PR 04-APR-2001; 2001US-281194P.
 PR 13-APR-2001; 2001US-283675P.
 PR 30-APR-2001; 2001US-287424P.
 PR 02-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288342P.
 PR 15-MAY-2001; 2001US-288528P.
 PR 16-MAY-2001; 2001US-291199P.
 PR 16-MAY-2001; 2001US-291099P.
 PR 30-MAY-2001; 2001US-291240P.
 PR 31-MAY-2001; 2001US-294485P.
 PR 31-MAY-2001; 2001US-294889P.
 PR 18-JUN-2001; 2001US-299027P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 19-JUN-2001; 2001US-299310P.
 PR 10-JUL-2001; 2001US-304354P.
 PR 31-JUL-2001; 2001US-309198P.
 PR 16-AUG-2001; 2001US-312903P.
 PR 10-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 27-SEP-2001; 2001US-325681P.
 PR 18-OCT-2001; 2001US-330380P.
 PR 31-OCT-2001; 2001US-335301P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-333184P.
 PR 14-NOV-2001; 2001US-333272P.
 PR 21-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 07-MAR-2002; 2002US-0092900.
 XX
 PA (CURA-) CURAGEN CORP.

XX Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L;
 PI Zechusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Palturanian M, Gangolli E, Vernet CAM, Guo X, Tchermey V;
 PI Fernandes ER, Casman SJ, Malyanar UM, Gerlach V, Liu Y;
 PI Andersen D, Spaderna SK, Catterton E, Burgess C, Lette M, Zhong H;
 PI Alsdbrook JP, Lepley DM, Rieger DK;
 XX
 DR WPI; 2002-723332/78.
 DR N-PDB; ABX97050.

PT NOXV polypeptides and polynucleotides, useful for preventing or
 PT treating a disorder associated with aberrant NOXV expression or
 PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
 PT bronchial asthma -
 XX
 PS Claim 1; Page 170; 1103pp; English.

XX This invention describes novel human NOXV polypeptides which have
 CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and
 CC hypotensive activity. Pharmaceutical compositions comprising the NOXV
 CC proteins or nucleic acid molecules or NOXV antibodies are useful for
 CC preventing or treating a disorder associated with aberrant NOXV
 CC expression or activity e.g. cancer, hypertension, atherosclerosis,
 CC cardiomyopathy or bronchial asthma. The products of the invention can
 CC be used for gene therapy or in a vaccine. ABU5041-ABU5218 represent
 CC the NOXV polypeptides encoded by ABX97008-ABX97185.

XX Sequence 1423 AA;

Query Match 95.1%; Score 7383.5; DB 23; Length 1423;
 Best Local Similarity 96.4%; Pred. No. 0;
 Matches 1408; Conservative 5; Mismatches 8; Indels 39; Gaps 9;

QY 1 MGLRPGIFLELLLELLLGQGTPOIHTSPKSTLEGQLMPEITKUSPEPCNPNKLGVMPSK 60
 DB 1 MGLRPGIFLELLLELLLGQGTPOIHTSPKSTLEGQLMPEITKUSPEPCNPNKLGVMPSK 40
 QY 61 SVHSKRPDIKFAAIGNLEIPDPGTGDEKQDWTERPOQVMGVMTVLSDIIRYSPS 120
 DB 41 -VHSKRPDIKFAAIGNLEIPDPGTGDEKQDWTERPOQVMGVMTVLSDIIRYSPS 97
 QY 121 VPMVCHTGRVIPHGAEDLWQAQELVNMKENQLDPOQPMKLNVEFSNASQCYLC 180
 DB 98 VPMVCHTGRVIPHGAEDLWQAQELVNMKENQLDPOQPMKLNVEFSNASQCYLC 156
 QY 181 PSAQONGLAAGVDELMGVLDYQOEVPFAFVNLSEVAEVSROYHTMTLPAPEPCN 240
 DB 157 PSAQONGLAAGVDELMGVLDYQOEVPFAFVNLSEVAEVSROYHTMTLPAPEPCN 216
 QY 241 CSEETTRAKVVMQMSYOEAMNSLASRYSQESFTVFPFPFYETTPSLHSBPRLOD 300
 DB 217 CSEETTRAKVVMQMSYOEAMNSLASRYSQESFTVFPFPFYETTPSLHSBPRLOD 272
 QY 301 STTLAMHLNRMMEPRAGEDEPLSVYGRPMKPSQESPTLFSYRNSNYLTRLOKPODKL 360
 DB 273 STTLAMHLNRMMEPRAGEDEPLSVYGRPMKPSQESPTLFSYRNSNYLTRLOKPODKL 332
 QY 361 EVREGAEIRCPDPSDPTVPTSVHRLKPADINVIGALGDSLTAONGAGSTPGNVLDVLTQ 420
 DB 333 EVREGAEIRCPDPSDPTVPTSVHRLKPADINVIGALGDSLTAONGAGSTPGNVLDVLTQ 392
 QY 421 YRGLSMVSGDENIGVTTLANILREBNSLKGFSVGTGKETSPNAFLNOAVAGRAEDL 480
 DB 393 YRGLSMVSGDENIGVTTLANILREBNSLKGFSVGTGKETSPNAFLNOAVAGRAEDL 450
 QY 481 PVQARLVLMKQDTRIHQEDPKITTLFIGNDLDCPDNDLVHSPONTNDIGRALDI 540
 DB 451 --QARLVLMKQDTRIHQEDPKITTLFIGNDLDCPDNDLVHSPONTNDIGRALDI 508
 QY 541 LHAEPRAFNLTVALEIVNLRELYOEKKYCPRMILRSICPVLKFPDDSTELATLIEF 600
 DB 509 LHAEPRAFNLTVALEIVNLRELYOEKKYCPRMILRSICPVLKFPDDSTELATLIEF 568
 QY 601 NKKEOETHQILBSGYDREDEFTVVVQPFENVMDPKTSEGLPDNSFFAPDCFFHSKKS 660
 DB 569 NKKEOETHQILBSGYDREDEFTVVVQPFENVMDPKTSEGLPDNSFFAPDCFFHSKKS 628
 QY 661 HSRPAASALNMNLEPVQOKTTRKFKENKINITCPNOVQ--PFLATYKNSQNGHGTWLP 719
 DB 629 HSRPAASALNMNLEPVQOKTTRKFKENKINITCPNOVQ--PFLATYKNSQNGHGTWLP 688
 QY 720 RAPSALHPTSVHARPADIOVVAALGDSLTAONGIGSKPDDLPDVTTOYRGLSYSAGDG 779
 DB 689 RAPSALHPTSVHARPADIOVVAALGDSLTAONGIGSKPDDLPDVTTOYRGLSYSAGDG 748
 QY 780 SLENVTTLPNILEFRNRLTGVAVGTGDANDTNAFLNOAVGAKAEDLMSQVOTLMQK 839
 DB 749 SLENVTTLPNILEFRNRLTGVAVGTGDANDTNAFLNOAVGAKAEDLMSQVOTLMQK 808
 QY 840 DHRVNFHEDEWKITVILIGSDLCYCTDSNLSAANFVDHLNADLVHREYPRVLVNL 899
 DB 809 DHRVNFHEDEWKITVILIGSDLCYCTDSNLSAANFVDHLNADLVHREYPRVLVNL 868
 QY 900 VDLNPLTMQVFLGPNDCPVQOASVLCVCTLLENSOELARLEAFSAYSSMRELV 959
 DB 869 VDLNPLTMQVFLGPNDCPVQOASVLCVCTLLENSOELARLEAFSAYSSMRELV 928
 QY 960 GSGRYTQEDSFVVLQPFQNIQLPVYLADLPDTSFFAPDCIHPNOKFHSQALRALMTNM 1019
 DB 929 GSGRYTQEDSFVVLQPFQNIQLPVYLADLPDTSFFAPDCIHPNOKFHSQALRALMTNM 988
 QY 1020 LEPGSKTETLIDRAEMPITCPTONEPFLRTPNNSNTYIKALIEWGSDFLCTEKWAS 1079
 DB 989 LEPGSKTETLIDRAEMPITCPTONEPFLRTPNNSNTYIKALIEWGSDFLCTEKWAS 1048

QY 1080 NSVPTSVHQLRPADIKVNAALGDSLTAVGARPNNSSDLPTSWRGLSMSIGDGNLETHT 1139
 Db 1049 NSVPTSVHQLRPADIKVNAALGDSLTVAVGARPNNSSDLPTSWRGLSMSIGDGNLETHT 1108
 QY 1140 TLNNILKKRPYLLGRTSTSTWEGTAGLVNAABEGARPADMAQMDLVERKNSP-DINLE 1198
 Db 1109 TLDDILKKRPYLLGRTSTSTWEGTAGLVNAABEGARPADMAQMDLVERKNSP-DINLE 1168
 QY 1199 KDKMLTLFIGNVDLGHYCNENPAHLATEYVQIQQALDILISELPRAFNVEVWELAS 1258
 Db 1169 KDKMLTLFIGNVDLGHYCNENPAHLATEYVQIQQALDILISELPRAFNVEVWELAS 1224
 QY 1259 LVQGGGKCAMLAANNCTCLRHSSQSLEKQELKKVWNLIQHGISSEFSYHQYTORDEPA 1318
 Db 1225 LVQGGGKCAMLAANNCTCLRHSSQSLEKQELKKVWNLIQHGISSEFSYHQYTORDEPA 1284
 QY 1319 VVVQPFQNTLTPLNERGDTLTFFPSHDCPHFSDRGAEVAIALMNNMLEPVGRKTTSSN 1378
 Db 1285 VVVQPFQNTLTPLNERGDTLTFFPSHDCPHFSDRGAEVAIALMNNMLEPVGRKTTSSN 1343
 QY 1379 FTHSRAKLKCPSPESPLYLTLRNSRLLPDQAEAPVLYWAVPVAAGVGLVGIIGTVW 1438
 Db 1344 FTHSRAKLKCPSPESPLYLTLRNSRLLPDQAEAPVLYWAVPVAAGVGLVGIIGTVW 1403
 QY 1439 RCRGGRREDPPMSLRTVAL 1458
 Db 1404 RCRGGRREDPPMSLRTVAL 1423

RESULT 5
 ABU65081
 ID ABU65081 standard; Protein; 1419 AA.
 XX

AC ABU65081;

DT 20-MAY-2003 (first entry)

DE Human NOV24a protein.

XX NOVA; cytosolic; cardiac; antiarteriosclerotic; antiasthmatic; cancer;
 KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 human.

XX Homo sapiens.

XX MO200272757-A2.

PD 19-SEP-2002.

XX 08-MAR-2002; 2002MO-US06908.

PR 08-MAR-2001; 2001US-274101P.
 PR 08-MAR-2001; 2001US-274194P.
 PR 08-MAR-2001; 2001US-274281P.
 PR 08-MAR-2001; 2001US-274322P.
 PR 09-MAR-2001; 2001US-274849P.
 PR 12-MAR-2001; 2001US-275235P.
 PR 13-MAR-2001; 2001US-275578P.
 PR 13-MAR-2001; 2001US-275579P.
 PR 13-MAR-2001; 2001US-275601P.
 PR 14-MAR-2001; 2001US-276000P.
 PR 16-MAR-2001; 2001US-276776P.
 PR 19-MAR-2001; 2001US-276994P.
 PR 20-MAR-2001; 2001US-277239P.
 PR 20-MAR-2001; 2001US-277321P.
 PR 20-MAR-2001; 2001US-277327P.
 PR 21-MAR-2001; 2001US-277791P.
 PR 22-MAR-2001; 2001US-277833P.
 PR 23-MAR-2001; 2001US-278152P.
 PR 26-MAR-2001; 2001US-278894P.
 PR 27-MAR-2001; 2001US-278999P.
 PR 27-MAR-2001; 2001US-279036P.
 PR 28-MAR-2001; 2001US-279344P.

PR 30-MAR-2001; 2001US-277338P.
 PR 30-MAR-2001; 2001US-279995P.
 PR 30-MAR-2001; 2001US-280233P.
 PR 02-APR-2001; 2001US-280802P.
 PR 02-APR-2001; 2001US-280822P.
 PR 02-APR-2001; 2001US-280900P.
 PR 04-APR-2001; 2001US-281194P.
 PR 13-APR-2001; 2001US-283675P.
 PR 30-APR-2001; 2001US-287442P.
 PR 02-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288342P.
 PR 03-MAY-2001; 2001US-288528P.
 PR 15-MAY-2001; 2001US-291190P.
 PR 16-MAY-2001; 2001US-291099P.
 PR 16-MAY-2001; 2001US-291240P.
 PR 30-MAY-2001; 2001US-294485P.
 PR 31-MAY-2001; 2001US-294889P.
 PR 31-MAY-2001; 2001US-294899P.
 PR 18-JUN-2001; 2001US-299027P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 19-JUN-2001; 2001US-299310P.
 PR 10-JUL-2001; 2001US-304354P.
 PR 31-JUL-2001; 2001US-309198P.
 PR 16-AUG-2001; 2001US-312903P.
 PR 10-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 27-SEP-2001; 2001US-325681P.
 PR 18-OCT-2001; 2001US-330380P.
 PR 31-OCT-2001; 2001US-335301P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-333184P.
 PR 14-NOV-2001; 2001US-333272P.
 PR 21-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 07-MAR-2002; 2002US-0092900.

(CURA-) CURAGEN CORP.

PA Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L;
 XX Zernhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Paturajan M, Gangolli E, Vernet CM, Guo X, Tcherven V;
 PI Fernandes ER, Casman SJ, Malvankar UM, Gerlach V, Liu Y;
 PI Anderson D, Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H;
 PI Alsobrook JP, Lepley DW, Rieger DK;

DR WPI; 2002-723332/78.
 DR N-PSDB; ABX97048.

XX NOVX polypeptides and polynucleotides, useful for preventing or
 PT treating a disorder associated with aberrant NOVX expression or
 PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
 PT bronchial asthma
 XX
 PS Claim 1; Page 168-169; 1103pp; English.

XX This invention describes novel human NOVX polypeptides which have
 CC cytostatic, cardiac, antiarteriosclerotic, antiasthmatic and
 CC hypotensive activity. Pharmaceutical compositions comprising the NOVX
 CC proteins or nucleic acid molecules or NOVX antibodies are useful for
 CC preventing or treating a disorder associated with aberrant NOVX
 CC expression or activity e.g. cancer, hypertension, atherosclerosis,
 CC cardiomyopathy or bronchial asthma. The products of the invention can
 CC be used for gene therapy or in a vaccine. ABU65041-ABU65218 represent
 CC the NOVX polypeptides encoded by ABX97008-ABX97185.

XX Sequence 1419 AA;

SO

DR WPI; 2002-426287/45.
 DR N-PSDB; AAD37410.
 XX
 PT New human phospholipase-like enzyme polypeptide useful for screening
 PT agents, and in the treatment of cancer, inflammation, diabetes,
 PT obesity, a central nervous system disorder, or a cardiovascular
 PT disorder
 XX
 PS Claim 25; Fig 2; 144pp; English.
 XX
 CC The present invention relates to novel human phospholipase-like enzymes
 CC and polynucleotides encoding such proteins. Sequences of the invention
 CC are useful for producing a medicament for modulating the activity of
 CC phospholipase in a disease such as cancer, inflammation, cardiovascular
 CC disorders, chronic obstructive pulmonary diseases, central nervous system
 CC (CNS) disorders such as brain injuries, cerebrovascular disease, dementia
 CC (Alzheimer's disease), Parkinson's disease, corticobasal degeneration,
 CC motor neuron disease, Pick's disease, Huntington's disease, Creutzfeldt
 CC Jacob dementia, schizophrenia with dementia, Korsakoff's psychosis,
 CC pain associated with CNS (e.g. epilepsy, failed back surgery syndrome,
 CC sciatica), multiple sclerosis, stroke, age associated memory impairment,
 CC allergic disease including asthma, allergic rhinitis (hay fever), atopic
 CC dermatitis, anaphylaxis and inflammation, cardiovascular disease, chronic
 CC obstructive pulmonary disease, acute respiratory distress syndrome, gout,
 CC diabetes, emphysema or obesity. They are also used for treating anorexia,
 CC overweight, cachexia, bulimia, hypertension, type-II diabetes, coronary
 CC artery disease, hyperlipidaemia, gall bladder disease, osteoarthritis,
 CC sleep apnoea and respiratory problems, cancer (e.g. breast, prostate,
 CC colon cancer), thrombolytic disease, reduced fertility, polycystic
 CC ovarian syndrome, complications of pregnancy, menstrual irregularity,
 CC hirsutism, stress incontinence and depression. The present sequence
 CC is human phospholipase like enzyme.
 CC
 XX
 SQ Sequence 1216 AA;
 Query Match 80.5%; Score 6253.5; DB 23; Length 1216;
 Best Local Similarity 97.1%; Pred. No. 0;
 Matches 1181; Conservative 7; Mismatches 9; Indels 19; Gaps 2;
 QY 259 EAMNSILASSRSYSEGSFTYVFPFPFYETTPSLHSEDPRLQDSTTLAHLMMNMEPAGE 318
 DB 1 EAMNSILASSRSYSEGSFTYVFPFPFYETTPSLHSEDPRLQDSTTLAHLMMNMEPAGE 60
 QY 319 KQEPILSVKGRPMKCSQESPYLFSYRNSNYLTRLQKPODKL-----EYRE 364
 DB 61 KQEPILSVKGRPMKCSQESPYLFSYRNSNYLTRLQKPODKL-----EYRE 120
 QY 365 GAETRCPPDPSPTVPSVHRLKPADINVGALGDSITAGAGSTPGNVLDVLTQYRGL 424
 DB 121 GAETRCPPDPSPTVPSVHRLKPADINVGALGDSITAGAGSTPGNVLDVLTQYRGL 180
 QY 425 SMSVGSDENIGVTTLANILREFNPSLKGFSVGTGKETSNAFLNQA VAGGRAEDLPVQA 484
 DB 181 SMSVGSDENIGVTTLANILREFNPSLKGFSVGTGKETSNAFLNQA VAGGRAEDLPVQA 240
 QY 485 RRLVDMKNDTRHHPQEDMKITLFTGNDLCPQNDLVHYSQNTDINIGKLDLTHAE 544
 DB 241 RRLVDMKNDTRHHPQEDMKITLFTGNDLCPQNDLVHYSQNTDINIGKLDLTHAE 300
 QY 545 VPRAFNLVTVLEIVNLRELJOEKYVCPMILRSICPCVLKFPDDNSTELATLIEFNKKE 604
 DB 301 VPRAFNLVTVLEIVNLRELJOEKYVCPMILRSICPCVLKFPDDNSTELATLIEFNKKE 360
 QY 605 QEXTHQLIESGRYDREDFTVVQVFPFENVDMKTSGLPDSNFFAPDCPHFSKSHSRA 664
 DB 361 QEXTHQLIESGRYDREDFTVVQVFPFENVDMKTSGLPDSNFFAPDCPHFSKSHSRA 420
 QY 665 ASALMMNMLEPVGQKTRHFKFNKITCPNOVQPFIRTK-----NSMOGHGTWLPQRD 719
 DB 421 ASALMMNMLEPVGQKTRHFKFNKITCPNOVQPFIRTK-----NSMOGHGTWLPQRD 480
 QY 720 RAPSALHPTSVHALRPADIQVVAALGDSLTPAGNGISGKPPDDLDPVTTYRGLSYSGDGS 779

DB 481 RAPSALHPTSVHALRPADIQVVAALGDSLTPAGNGISGKPPDDLDPVTTYRGLSYSGDGS 540
 QY 780 SLENTVTLNRIAREFNRLTGVAVGNDANDTNAFLNQA VPGAEDLMSQVOTLQOMK 839
 DB 541 SLENTVTLNRIAREFNRLTGVAVGNDANDTNAFLNQA VPGAEDLMSQVOTLQOMK 600
 QY 840 DHRVNFHEDMKVITLIGGSDLCDYCTDSNLSAANFVDHLNNAALDVHREPRVLVNL 899
 DB 601 DHRVNFHEDMKVITLIGGSDLCDYCTDSNLSAANFVDHLNNAALDVHREPRVLVNL 660
 QY 900 VDFLNPITKQVFLGNPDCKPVQASVLCNCVLTLENSQELARLEAFSAVSSMRELV 959
 DB 661 VDFLNPITKQVFLGNPDCKPVQASVLCNCVLTLENSQELARLEAFSAVSSMRELV 720
 QY 960 GSGRYTOTDPSVYVLPFQNTQLPVLADGLPOTSFPAPDCTHPNOKFHSQILRALMTNM 1019
 DB 721 GSGRYTOTDPSVYVLPFQNTQLPVLADGLPOTSFPAPDCTHPNOKFHSQILRALMTNM 780
 QY 1020 LEPGSKTETLDRAEMPTTCPTQNEPFLRTPPNSVYTPYIKPAIENWGSDFLCTEWEKAS 1079
 DB 781 LEPGSKTETLDRAEMPTTCPTQNEPFLRTPPNSVYTPYIKPAIENWGSDFLCTEWEKAS 840
 QY 1080 NSVPTSVMQLRPADIKVVAALGDSLTPAVGARPNSSDLPTSNGLSWSIGDGNLETHT 1139
 DB 841 NSVPTSVMQLRPADIKVVAALGDSLTPAVGARPNSSDLPTSNGLSWSIGDGNLETHT 900
 QY 1140 TLPNILKKNPYLIGSTSTWECTAGLNAABEABARBDPAQMDLVERKNSPDITLLEK 1199
 DB 901 TLPNILKKNPYLIGSTSTWECTAGLNAABEABARBDPAQMDLVERKNSPDITLLEK 960
 QY 1200 DMKLVTLFTIGVNDLCHYCNENPEALHATEYVOHIOALDILSEELPRAFVVVEVMEIASL 1259
 DB 961 DMKLVTLFTIGVNDLCHYCNENPEALHATEYVOHIOALDILSEELPRAFVVVEVMEIASL 1020
 QY 1260 YQGGGKCAMLAAQNNCTCLRHSQSLERQELKKNVNLQHGISSFSYHQYTORQEDPAV 1319
 DB 1021 YQGGGKCAMLAAQNNCTCLRHSQSLERQELKKNVNLQHGISSFSYHQYTORQEDPAV 1080
 QY 1320 VVQPFQNTLTPLANEGDPTDLTFPSQDCPHFSRGAENALIAMNNMLEPVGKRTTSNNF 1379
 DB 1081 VVQPFQNTLTPLANEGDPTDLTFPSQDCPHFSRGAENALIAMNNMLEPVGKRTTSNNF 1140
 QY 1380 TTSRAKLKCPSPSPYLYTLRNSRLPDDQAEABEVLVYNAVPVAGVGLVVGIIIGTVNR 1439
 DB 1141 TTSRAKLKCPSPSPYLYTLRNSRLPDDQAEABEVLVYNAVPVAGVGLVVGIIIGTVNR 1200
 QY 1440 CRRGGRREDPPMSLRT 1455
 DB 1201 CRRGGRREDPPMSLRT 1216
 RESULT 7
 AAM30751
 ID AAM30751 standard; Protein; 1450 AA.
 XX
 AC AAM30751;
 XX
 DT 28-JAN-1998 (first entry)
 XX
 DE Rat phospholipase-B/lipase.
 XX
 KW phospholipase B; lipase; supplement; pancreatic phospholipase; reagent;
 XX screening; rat.
 XX
 OS Rattus rattus.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..30 /label= signal_peptide
 FT Protein 31 /label= mature_protein
 FT Region 43..652 /label= repeat_region_1

```

FT Region 367..712 /label= repeat_region_2
FT Region 714..1059 /label= repeat_region_3
FT Region 1070..1408 /label= repeat_region_4
FT Domain 1421..1443 /label= transmembrane_binding_domain
PN JP09248190-A.
PD 22-SEP-1997.
PF 15-MAR-1996; 96JP-0086022.
PR 15-MAR-1996; 96JP-0086022.
XX (TOJO/) TOJO H.
XX (TOFU) TONEN CORP.
XX MPI; 1997-520744/48.
XX N-PSDB; AAT51874.
XX DNA encoding protein with phospholipase B and lipase activity -
PT useful as supplemant to pancreatic phospholipase
XX Claim 1; Pages 7-12; 16pp; Japanese.
XX The present sequence is a protein which has both phospholipase B and
CC lipase activity. The claimed protein comprises at least amino acids
CC 367-712 of this sequence. The phospholipase is useful as a supplemant to
CC pancreatic phospholipase and as a reagent for the determination of
CC phospholipase B/lipase.
SQ Sequence 1450 AA:
Query Match 68.0%; Score 5279.5; DB 18; Length 1450;
Best Local Similarity 70.0%; Pred. No. 0;
Matches 1012; Conservative 161; Mismatches 263; Indels 9; Gaps 6;
QY 5 PGIFL-LELILLILGOGTPOIHSPRKSTLEGLMPETLNKSPPCNPKLGVNMPKSV 62
DB 5 PGVSLVGLILLILLGQSPQIHGSSGENTSQPOQVFTLNKFSFCKPKLESLVSKSV 64
QY 63 HSLKSPDIDKVAIIGNLEIETPPDGTG--DLEK-QDWTERPOQCMGMVTLSDIIRFSP 119
DB 65 HSLRPSDIDKVAIIGNLEIETPPAPGSGVNNKPKQSLSELDQNCIGIMTSLDIIRFNP 124
QY 120 SVMPVPCHTGKRYIAPHDG-AEDLMIQAEIIVRNKKNLQDFOPDWKLINVPFNSAQCY 178
DB 125 SVLMPTCSPKGTAGHTTIAEDLMIQAKELVRHKKNPEIDFEKDWKLITVLFSNTSQCH 184
QY 179 LCPSAQONGIAGVDELMLGVLDYLQOEVPRAPVNLVLSVAEVSRYHGTWLSAPAP 238
DB 185 LCSSDQCKRILMKHMEVLSGLVDYLHREVPRAPVNLVLSVAEVSRYHGTWLSAPAP 244
QY 239 CNGSEETRLAKVYMQMSYQEAANSLAASRYSQESFTVFPQFETTPSLHSEPR 298
DB 245 CKSESEETRLAKVYMQMSYQEAANSLAASRYSQESFTVFPQFETTPSLHSEPR 302
QY 299 QDSSTLLAHLMMNMMEPAGKEDEPLSYVHGRPMKCPQESPYLFSYNSNYLQRLQKPD 358
DB 303 QDSSTLLAHLMMNMMEPAGKEDEPLSYVHGRPMKCPQESPYLFSYNSNYLQRLQKPD 362
QY 359 KLEVRGAEIRCPDKSDPTVPTSVHRLKPADINIVIGALDLSLTAAGASTPGNVLDVL 418
DB 363 KFMKEGTCTCPDKSDPTVPTSVHRLKPADINIVIGALDLSLTAAGASTPGNVLDVL 422
QY 419 TQYRGISWSVSGDENIGTVTTLANIIEFNPSLKGESVGTGKETSPVAFLQAVAGRAE 478
DB 423 TQYRGISWSVSGDETITETVTTLANIIEFNPSLKGESVGTGKETSPVAFLQAVAGRAE 482
QY 479 DLVQARLVLDMKNDTRIHQEDWKITITLFIGNDLQDPCNDLVHYSQNPFTONICKAL 538

```

```

DB 483 GLAAQAKKLVSIMKDDKTIHQEDWKIITVFIGNDLQDSCNNLAFSPQTFDINKITLAL 542
QY 539 DILHAEPAPVNLVTLVLEIVNLRELYQEKTYCCPMILIRSLCPCLAKKDDNSTELATLI 598
DB 543 DILHAEPAPVNLVTLVLEIVNLRELYQEKTYCCPMILIRSLCPCLAKKDDNSTELATLI 602
QY 599 EFNKPKQEKTHOLIESGRYDTRDPFTVVOVQPFENVMPKTSSEGLPDNSFFAPDCFPSS 658
DB 603 EFNKPKQEKTHOLIESGRYDTRDPFTVVOVQPFENVMPKTSSEGLPDNSFFAPDCFPSS 662
QY 659 KSHSRASALAMNNMIEPVQOKTTRHKKENKINITCPNOVQPLRTYKSMQGHGTLPCR 718
DB 663 KTHARSALAMNNMIEPVQOKTTRHKKENKINITCPNOVQPLRTYKSMQGHGTLPCR 722
QY 719 DAPRSALHPTSVHAAIRPADIOVVAALGDSLTGNGISGRPDLPVTTQYRGISYAGGD 778
DB 723 EKAPASAPPTSVHAAIRPADIOVVAALGDSLTGNGISGRPDLPVTTQYRGISYAGGD 782
QY 779 GSLENVTTLPILREFNRLVTGYAVGTGDPANDTNAFLNQAIPGAKAEDLMSQVOTLMQK 838
DB 783 KFLENVTTLPILREFNRLVTGYAVGTGDPANDTNAFLNQAIPGAKAEDLMSQVOTLMQK 842
QY 839 KODHRVNFEDMKVITVLIGSDLCDCYDSMLYSANFVHLRNALDYLHREVPRLVN 898
DB 843 KODTRVNFQDMKVITVMIIGASDLCDPCKDSNRYSANFSDHLRNALDYLHREVPRLVN 902
QY 899 LVDFLNPITMROVFLGNPKCPVQOASVYCNCVLTIRENSQELARLEASRAASRSMREL 958
DB 903 LVDFLNPITMROVFLGNPKCPVQOASVYCNCVLTIRENSQELARLEASRAASRSMREL 962
QY 959 VSGGRYDTEDEFSVVLQPFQONQLPVLADGLPDTFFAPDCIHPQCKHFSQALARMY 1018
DB 963 VSGGRYDTEDEFSVVLQPFQONQLPVLADGLPDTFFAPDCIHPQCKHFSQALARMY 1022
QY 1019 MLEPLGSKTETDILRAEMBITCPTONEPFLRTPRNSNYTYPIKPAIENNGSDFLCTEMKA 1078
DB 1023 MLEPLGSKTETDILRAEMBITCPTONEPFLRTPRNSNYTYPIKPAIENNGSDFLCTEMKA 1082
QY 1079 SNSVPTSVQOLRADITKVVAALGDSLTGAVGARPNNSDLPFSWRGLSWSIGDGLTETH 1138
DB 1083 SNSVPTSVQOLRADITKVVAALGDSLTGAVGARPNNSDLPFSWRGLSWSIGDGLTETH 1142
QY 1139 TILPNTILKGFNPYLDEFSTSTWEGTAGLVNVAEGARPMAPQAWDLVERMKSPDINLE 1198
DB 1143 TILPNTILKGFNPYLDEFSTSTWEGTAGLVNVAEGARPMAPQAWDLVERMKSPDINLE 1202
QY 1199 KDMKLTLEFIGVNDLCHYCENPEAHILATEYVQHIQALDILSEELPRAPVNVVEMELAS 1258
DB 1203 KDMKLTLEFIGVNDLCHYCENPEAHILATEYVQHIQALDILSEELPRAPVNVVEMELAS 1262
QY 1259 LYOGGQCKAM-LAANNCTCLRHSSQSLKQELKRVNNMLQGISFSYWHYOTQREDP 1317
DB 1263 LYOGGQCKAM-LAANNCTCLRHSSQSLKQELKRVNNMLQGISFSYWHYOTQREDP 1322
QY 1318 AVYVQPFQONTLPILNERGDTDLTPFSEDCPFHSDRGHAMAIALANNMIEPVGRKTTSN 1377
DB 1323 AVYVQPFQONTLPILNERGDTDLTPFSEDCPFHSDRGHAMAIALANNMIEPVGRKTTSN 1382
QY 1378 NFTHSRAKLKCPSPESPYLTYLNSRLLPQAEABEVLVYAVVAVAGVLVVGIIIGTVV 1437
DB 1383 NFTHSRAKLKCPSPESPYLTYLNSRLLPQAEABEVLVYAVVAVAGVLVVGIIIGTVV 1442
QY 1438 WRCCR 1442
DB 1443 WRCCR 1447

```

RESULT 8
 AAE34448 standard; Protein; 969 AA.
 ID AAE34448
 XX AAE34448;
 AC AAE34448;

QY 601 NKKFQEKHTOLIBSGRYDTEDEFTVVVQPFENVMPKTSGLPDNSFFAPDCFFHSKS 660
 Db 601 NKKFQEKHTOLIBSGRYDTEDEFTVVVQPFENVMPKTSGLPDNSFFAPDCFFHSKS 660
 QY 661 HSRPAASALNNMMLPEVQCKTTRHKFEENKINITCENOVQPLRTYKNSMOGHGWTLPQRDR 720
 Db 661 HSRPAASALNNMMLPEVQCKTTRHKFEENKINITCENOVQPLRTYKNSMOGHGWTLPQRDR 720
 QY 721 APSALHTSVHALRPADIQVVAALGDSLTTAGNGISGRPDLPVTTQYRGISYAGSDGS 780
 Db 721 APSALHTSVHALRPADIQVVAALGDSLTTAGNGISGRPDLPVTTQYRGISYAGSDGS 780
 QY 781 LENVTTLPNLTRENNLTGVAAGTGDANDTNAFINQAVGAKAEDLMSQVQTLMOQKMD 840
 Db 781 LENVTTLPNLTRENNLTGVAAGTGDANDTNAFINQAVGAKAEDLMSQVQTLMOQKMD 840
 QY 841 DHRVNFHEDMKVITVLLIGSDLCDCYCTDSNLYSANFVDHLRNALDYLAREVPRLVNLV 900
 Db 841 DHRVNFHEDMKVITVLLIGSDLCDCYCTDSNLYSANFVDHLRNALDYLAREVPRLVNLV 900
 QY 901 DFLNPTIMROVFLGNPDKCPVQQA 924
 Db 901 DFLNPTIMROVFLGNPDKCPVQQA 924

RESULT 9
 ID ABP53556 standard; Protein: 472 AA.
 AC ABP53556;
 DT 16-DEC-2002 (first entry)
 XX Human phospholipase protein SEQ ID NO:2.
 KW Human; phospholipase; enzyme; chromosome 2.
 OS Homo sapiens.
 PN NO200262977-A2.
 PD 15-AUG-2002.
 PF 28-JAN-2002; 2002WO-US02302.
 PR 08-FEB-2001; 2001US-0778961.
 XX (PEKE) PE CORP NY.
 PA Van C, Ketchum KA, Di Francesco V, Beasley EM;
 PI WPI; 2002-682698/73.
 DR N-PSDB; ABQ82234.
 XX New human phospholipase proteins, useful for the development of human
 PT therapeutics and diagnostic compositions, drug screening assays, tissue
 PT typing and pharmacogenomic analysis
 PS Claim 1; Fig 2A; 95bp; English.

The present sequence represents a human phospholipase protein (I) located on chromosome 2. (I) can be used for identifying agents that modulate its function or activity where the agent is useful for treating a disease or condition mediated by a the human phospholipase protein. (I) peptides can be used in substantial and specific assays related to functional information of the peptide sequences, to raise antibodies or to elicit immune response, as reagents in assays that determine the levels of protein in biological fluids, and as markers for tissues where the corresponding protein is expressed. Nucleotide sequences encoding (I) can be used as probes, primers and chemical intermediates in biological assays, for constructing recombinant vectors, and expressing antigenic portions of the protein. (I) and nucleic acid molecules encoding it can be used in the identification of therapeutic proteins and may serve as

CC models or targets for the development of human therapeutic agents that
 CC modulate phospholipase activity in cells and tissues that express the
 CC phospholipase, such as in kidney, blood, lung, brain glioblastoma,
 CC prostate, colon or leukocytes.
 XX Sequence 472 AA;
 SQ

Query Match 32.1%; Score 2492; DB 23; Length 472;
 Best local Similarity 93.7%; Pred. No. 2e-216;
 Matches 472; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 955 MRELVSGRYDTEDEFSVVLQPFQNIQLPVLADGLPDTSFAPDCIHNPQKHSQLARA 1014
 Db 1 MRELVSGRYDTEDEFSVVLQPFQNIQLPVLADGLPDTSFAPDCIHNPQKHSQLARA 1014
 QY 1015 LMTNMLEPLGSKTETDLRAEMPTCTPTONEPRLTPRNSNTYTPKPALENMGSDLCT 1074
 Db 34 -----LEPLGSKTETDLRAEMPTCTPTONEPRLTPRNSNTYTPKPALENMGSDLCT 88
 QY 1075 EMKASNSVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSDLPSTMRGLSMSIGSDGN 1134
 Db 89 EMKASNSVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSDLPSTMRGLSMSIGSDGN 1134
 QY 1135 LETHITLPLNLTKEFNPLLGFTSTWEGTAGLNVALEGARDPMDPAQMDLVERMKNSPD 1194
 Db 149 LETHITLPLNLTKEFNPLLGFTSTWEGTAGLNVALEGARDPMDPAQMDLVERMKNSPD 1194
 QY 1195 INLEKMKVLTFLTGNDLCHYCEPNAHLATRYVQHIOQALDILSEELPRAVNVEVM 1254
 Db 209 INLEKMKVLTFLTGNDLCHYCEPNAHLATRYVQHIOQALDILSEELPRAVNVEVM 1254
 QY 1255 ELASIVOGGKCAMLAQNNCTCLRHSQSLEKQELKYNMNIQHGISSFSYWHQYTOR 1314
 Db 269 ELASIVOGGKCAMLAQNNCTCLRHSQSLEKQELKYNMNIQHGISSFSYWHQYTOR 1314
 QY 1315 EDFAVVVOFPFQNTLPLNLRGDTDLTFESGDCFHFSDRGHAMAILANNMMLPEVGRKT 1374
 Db 329 EDFAVVVOFPFQNTLPLNLRGDTDLTFESGDCFHFSDRGHAMAILANNMMLPEVGRKT 1374
 QY 1375 TSNNFTHSRAKLKCPSESPYLLTLRNSRLLPQAEAPRVLYMAVPAAGVLVGGIIG 1434
 Db 389 TSNNFTHSRAKLKCPSESPYLLTLRNSRLLPQAEAPRVLYMAVPAAGVLVGGIIG 1434
 QY 1435 TVVMRCRGRGRREDPPMSLRTVAL 1458
 Db 449 TVVMRCRGRGRREDPPMSLRTVAL 1458

RESULT 10
 ID ABU65082 standard; Protein: 310 AA.
 AC ABU65082;
 DT 20-MAY-2003 (first entry)
 XX Human NOV24b protein.
 KW NOVX; cytosolic; cardiac; antiarteriosclerotic; antiasthmatic; cancer;
 KM hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 KM human.
 OS Homo sapiens.
 PN NO200272757-A2.
 PD 19-SEP-2002.
 PF 08-MAR-2002; 2002WO-US06908.
 PR 08-MAR-2001; 2001US-274101P.
 PR 08-MAR-2001; 2001US-274194P.
 PR 08-MAR-2001; 2001US-274281P.

PR 08-MAR-2001; 2001US-274322P.
 PR 09-MAR-2001; 2001US-274849P.
 PR 12-MAR-2001; 2001US-275235P.
 PR 13-MAR-2001; 2001US-275578P.
 PR 13-MAR-2001; 2001US-275579P.
 PR 13-MAR-2001; 2001US-275601P.
 PR 14-MAR-2001; 2001US-276000P.
 PR 16-MAR-2001; 2001US-276776P.
 PR 19-MAR-2001; 2001US-276994P.
 PR 20-MAR-2001; 2001US-277239P.
 PR 20-MAR-2001; 2001US-277321P.
 PR 20-MAR-2001; 2001US-277327P.
 PR 21-MAR-2001; 2001US-277791P.
 PR 22-MAR-2001; 2001US-277833P.
 PR 23-MAR-2001; 2001US-278152P.
 PR 26-MAR-2001; 2001US-278894P.
 PR 27-MAR-2001; 2001US-278999P.
 PR 27-MAR-2001; 2001US-279036P.
 PR 28-MAR-2001; 2001US-279344P.
 PR 30-MAR-2001; 2001US-279995P.
 PR 30-MAR-2001; 2001US-280233P.
 PR 02-APR-2001; 2001US-280802P.
 PR 02-APR-2001; 2001US-280822P.
 PR 02-APR-2001; 2001US-280900P.
 PR 04-APR-2001; 2001US-281194P.
 PR 13-APR-2001; 2001US-283675P.
 PR 30-APR-2001; 2001US-287424P.
 PR 02-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288342P.
 PR 03-MAY-2001; 2001US-288528P.
 PR 15-MAY-2001; 2001US-291190P.
 PR 16-MAY-2001; 2001US-291099P.
 PR 16-MAY-2001; 2001US-291240P.
 PR 30-MAY-2001; 2001US-294485P.
 PR 31-MAY-2001; 2001US-294889P.
 PR 31-MAY-2001; 2001US-294999P.
 PR 18-JUN-2001; 2001US-299027P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 19-JUN-2001; 2001US-299310P.
 PR 10-JUL-2001; 2001US-304354P.
 PR 31-JUL-2001; 2001US-309198P.
 PR 16-AUG-2001; 2001US-312903P.
 PR 10-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 27-SEP-2001; 2001US-325681P.
 PR 18-OCT-2001; 2001US-330380P.
 PR 31-OCT-2001; 2001US-335301P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-333184P.
 PR 14-NOV-2001; 2001US-333272P.
 PR 21-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 07-MAR-2002; 2002US-0092900.
 XX
 RA (CURA-) CURAGEN CORP.
 XX
 PI Padigaru M, Szytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L,
 PI Zernusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Paturajan M, Gangoli E, Vernet CAM, Guo X, Tcherev V;
 PI Fernandez ER, Casman SJ, Malvankar UM, Gerlach V, Liu Y;
 PI Anderson D, Spaderna SK, Catterton E, Burgess C, Lette M, Zhong H;
 PI Alsdorook JF, Lepley DM, Rieger DK;
 XX
 DR MPI; 2002-723332/78.
 DR N-PSDB; ABX97049.
 XX

PT NOXV polypeptides and polynucleotides, useful for preventing or
 PT treating a disorder associated with aberrant NOXV expression or
 PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
 PT bronchial asthma
 PS Claim 1; Page 169; 1103pp; English.
 XX
 CC This invention describes novel human NOXV polypeptides which have
 CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and
 CC hypotensive activity. Pharmaceutical compositions comprising the NOXV
 CC proteins or nucleic acid molecules or NOXV antibodies are useful for
 CC preventing or treating a disorder associated with aberrant NOXV
 CC expression or activity e.g., cancer, hypertension, atherosclerosis,
 CC cardiomyopathy or bronchial asthma. The products of the invention can
 CC be used for gene therapy or in a vaccine. ABUS5041-ABUS5218 represent
 CC the NOXV polypeptides encoded by ABX97008-ABX97185.
 XX
 SQ Sequence 310 AA;
 Query Match 19.6%; Score 1525; DB 23; Length 310;
 Best Local Similarity 97.3%; Pred. No. 5.2e-129;
 Matches 285; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 491 MKNDTRIHFOEDWKITLFIQGNLDFCNDLVHVSPOFTDNGALDILHAEPFAFV 550
 DB 1 MKNDTRIHFOEDWKITLFIQGNLDFCNDLVHVSPOFTDNGALDILHAEPFAFV 60
 QY 551 NLVTVEIYNLRELVOEKYVCPRMILRSJLPCVLKFDNDSTELATLIEFNKKFOEKTQ 610
 DB 61 NLVTVEIYNLRELVOEKYVCPRMILRSJLPCVLKFDNDSTELATLIEFNKKFOEKTQ 120
 QY 611 LTESGYDRDEPTVVVOPPEFVNDPKTSEGLPDNSFPAPDCFFRSKSHSAALMN 670
 DB 121 LIESGYDRDEPTVVVOPPEFVNDPKTSEGLPDNSFPAPDCFFRSKSHSAALMN 180
 QY 671 NMLEPVGQKTRHKFEFNKINITCPNOVOPFLRTYKNSMOGHGIMLPCRPAPSAALHPTSV 730
 DB 181 NMLEPVGQKTRHKFEFNKINITCPNOVOPFLRTYKNSMOGHGIMLPCRPAPSAALHPTSV 240
 QY 731 HALRPADIVVAALGDSLTAAGNGISGSKPDDLPTVTYQRLSYSGAGDSLEN 783
 DB 241 HALRPADIVVAALGDSLTAAGNGISGSKPDDLPTVTYQRLSYSGAGDSLEN 293
 RESULT 11
 ID ABB11053 standard; peptide; 267 AA.
 XX
 AC ABB11053;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human phospholipase B homologue, SEQ ID NO:1423.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; chromolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotrophic; cardiant; virocid; antibacterial;
 KW antifungal; vulnerary; antidiabetic.
 XX
 OS Homo sapiens.
 XX
 PN WO200157188-A2.
 XX
 PD 09-AUG-2001.

XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
XX WPI, 2001-463232/52.
XX
XX Nucleic acids encoding 973 human testicular antigen polypeptides,
PT useful for preventing, diagnosing and/or treating testicular cancer
XX
PS Claim 11; SEQ ID NO 1495; 766bp; English.
XX
XX The present invention provides the protein and coding sequences of 973
CC human testicular antigens, and fragments of their genomic sequences. The
CC sequences can be used in the treatment of cardiovascular, urinary system,
CC reproductive system, immune, respiratory, neurological and
CC gastrointestinal disorders, infections, and particularly cancer,
CC especially testicular cancers. The present sequence is a protein of the
CC invention.
XX
SQ Sequence 148 AA;

Query Match 7.2%; Score 560; DB 22; Length 146;
Best Local Similarity 77.9%; Pred. No. 4.9e-42;
Matches 113; Conservative 11; Mismatches 17; Indels 4; Gaps 3;

QY 1013 RAIVTMLEPLGSKETETLDRAEMPTCPQNEPFLTPNNSVYTPKPAIENMGSDPL 1072
DB 4 QSLMTNNLEPLGSKETETLDRAEMPTCPQNEPFLTPNNSVYTPKPAIENMGSDPL 63
QY 1073 CTEWKASNSVPTSVHQLRPADIKVVALGDSLTTAVGAPRNSSDPTSWRG-LSWSIGG 1131
DB 64 CTEWKASNSVPTSVHQLRPADIKVVALGDSLTTAVGAPRNSSDPTSWRG-LSWSIGG 123
QY 1132 DGNL-ETHHTLPNI--LKKENPYLL 1153
DB 124 MGNLGDSHHTAQHSEEVQPLPFWLL 148

RESULT 15
AAM95420
ID AAM95420 standard; Protein; 148 AA.
XX
AC AAM95420;
XX
DT 21-NOV-2001 (first entry)
XX
XX Human reproductive system related antigen SEQ ID NO: 4078.
DE Human reproductive system related antigen; reproductive system disorder;
XX
XX Human; reproductive system related antigen; reproductive system disorder;
KM cancer; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01339.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225457.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 05-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.

XX 01-AUG-2002.
 PD 22-JAN-2002; 2002MO-US01715.
 XX 24-JAN-2001; 2001US-264049P.
 XX (LEXI-) LEXICON GENETICS INC.
 PA Yu X, Miranda M, Turner CA;
 PI WPI; 2002-599797/64.
 DR P-PSDB; ABB09555.
 XX Polynucleotides encoding human lipases that are structurally related to
 PT animal lipases, particularly phospholipase B, useful for drug
 PT screening, diagnosis and in gene therapy of biological disorders
 PS Claim 1; Page 36-37; 44pp; English.
 XX The invention relates to a novel human lipase (NHL; ABB09555, ABB09556)
 CC and to nucleic acids encoding it (ABQ77623, ABQ77624). The NHL has
 CC structural similarity with animal lipases, particularly phospholipase B.
 CC Polynucleotides encoding NHL were obtained using human genomic sequences
 CC in conjunction with human thyroid and brain cDNAs. The NHL gene is
 CC located on chromosome 2, and contains a C/T polymorphism at position
 CC 3953 of the open reading frame (ORF), resulting in an Ala/Val
 CC substitution at position 1318 in the protein. NHL nucleotides and
 CC proteins are useful for treating disorders such as inflammatory or
 CC proliferative disease, infectious disease, clotting disorders, and
 CC cancer. They can also be used in screening for compounds useful in
 CC the treatment of mental, biological or medical disorders, as diagnostic
 CC reagents, in clinical trial monitoring and in cosmetic and nutritional
 CC applications. NHL nucleotides can additionally be used in the detection
 CC of disease-associated mutations, in the analysis of gene expression, for
 CC the recombinant expression of NHL, to generate transgenic animals, in
 CC gene therapy, and as part of ribozyme and/or triple helix sequences
 CC useful in the modulation of NHL gene expression. The present sequence
 CC represents cDNA encoding the Ala 1318 variant of NHL.
 XX
 SQ Sequence 4377 BP; 1085 A; 1217 C; 1155 G; 920 T; 0 other;
 Query Match 100.0%; Score 4377; DB 24; Length 4377;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 TTGGAATTCAGGCTCAAGAACTGTGAGAAACATGAAAGAACTGTCAACTT 480
 DB 421 TTGGAATTCAGGCTCAAGAACTGTGAGAAACATGAAAGAACTGTCAACTT 480
 QY 481 CAATTGACTGAAAGCTCATCAATGTGTCTTGAGTAATGCAAGCCAGTGTACTGTGC 540
 DB 481 CAATTGACTGAAAGCTCATCAATGTGTCTTGAGTAATGCAAGCCAGTGTACTGTGC 540
 QY 541 CCTCTGCTCAACAGAAATGGGCTTGGCGGGCGGCTGTGATAGCTGATGGGGTCTG 600
 DB 541 CCTCTGCTCAACAGAAATGGGCTTGGCGGGCGGCTGTGATAGCTGATGGGGTCTG 600
 QY 601 GACTACCTGACAGAGAGGTCCCAAGACATTTGTAAACCTGTGTGACCTCTGAGATT 660
 DB 601 GACTACCTGACAGAGAGGTCCCAAGACATTTGTAAACCTGTGTGACCTCTGAGATT 660
 QY 661 GCAGAGGCTCTGTGTAGTATCAAGGCACTTGGCTCAGCCCTGCAACCAAGCCCTGTAT 720
 DB 661 GCAGAGGCTCTGTGTAGTATCAAGGCACTTGGCTCAGCCCTGCAACCAAGCCCTGTAT 720
 QY 721 TGCTCAGAGAGAACACCCGCTGGCCAGAGTGTATGCAAGTGTATTATCAGAAAGCC 780
 DB 721 TGCTCAGAGAGAACACCCGCTGGCCAGAGTGTATGCAAGTGTATTATCAGAAAGCC 780
 QY 781 TGGAAACGCTCCGCGCTCAGACAGGTACAGTGAAGAGAGTCTTCAACGTTGTTTC 840
 DB 781 TGGAAACGCTCCGCGCTCAGACAGGTACAGTGAAGAGAGTCTTCAACGTTGTTTC 840
 QY 841 CAGCTTTCTTATGAGAACACCCCATCTCTACCTGAGAGACCCCGCACTTCAGAGAT 900
 DB 841 CAGCTTTCTTATGAGAACACCCCATCTCTACCTGAGAGACCCCGCACTTCAGAGAT 900
 QY 901 TCTAACAGCTGGCTGGCATCTCTGGAATAGATATGGAAGCCAGAGAGAGAAAGAT 960
 DB 901 TCTAACAGCTGGCTGGCATCTCTGGAATAGATATGGAAGCCAGAGAGAGAAAGAT 960
 QY 961 GAGCAATGAGTAAACACAGGAGGCAATGAAAGTCCCTCTCAGAGAGAGCCCTAT 1020
 DB 961 GAGCAATGAGTAAACACAGGAGGCAATGAAAGTCCCTCTCAGAGAGAGCCCTAT 1020
 QY 1021 CTGTTCAAGTACAGAAACACCAACTACCTGACAGACTGTGAGAAACCCCAAGACACTT 1080
 DB 1021 CTGTTCAAGTACAGAAACACCAACTACCTGACAGACTGTGAGAAACCCCAAGACACTT 1080
 QY 1081 GAGGTAAAGAAAGAGCGGAATCAATGATCTCTGACAAAGACCCCTCCGATACGTTCC 1140
 DB 1081 GAGGTAAAGAAAGAGCGGAATCAATGATCTCTGACAAAGACCCCTCCGATACGTTCC 1140
 QY 1141 ACCTCAGTTCATAGGCTGAAGCCGGCTGACATGAACTTAATTGGAGCCCTGGGTGACTCT 1200
 DB 1141 ACCTCAGTTCATAGGCTGAAGCCGGCTGACATGAACTTAATTGGAGCCCTGGGTGACTCT 1200
 QY 1201 CTCAGGCAAGGCAATGGGCGGGGTCCACACTGTGGAAACGTTTGAACGTTGACTCAG 1260
 DB 1201 CTCAGGCAAGGCAATGGGCGGGGTCCACACTGTGGAAACGTTTGAACGTTGACTCAG 1260
 QY 1261 TACCGAGGCTGTCTCTGAGCGTGGCGGAGATGAGAAATCGGCACTTTACCACTCTG 1320
 DB 1261 TACCGAGGCTGTCTCTGAGCGTGGCGGAGATGAGAAATCGGCACTTTACCACTCTG 1320
 QY 1321 GCGAACAATCTCCGGGAATTCAACCTTCCCTGAAGGGCTTCTCTGTGGCACTGGGAAA 1380
 DB 1321 GCGAACAATCTCCGGGAATTCAACCTTCCCTGAAGGGCTTCTCTGTGGCACTGGGAAA 1380
 QY 1381 GAAACAGTCTATATGCTTTTAAACAGAGCTGTGGCAGAGAGCCAGCTGAGATCTA 1440
 DB 1381 GAAACAGTCTATATGCTTTTAAACAGAGCTGTGGCAGAGAGCCAGCTGAGATCTA 1440
 QY 1441 CTTGTCCAGGCGCAGAGAGCTGTGTGACCTGTATGAAGATGACACAGAGATACACTTTCAG 1500
 DB 1441 CTTGTCCAGGCGCAGAGAGCTGTGTGACCTGTATGAAGATGACACAGAGATACACTTTCAG 1500

QY 1501 GAAGACTGGAAGATAAACCTGTTTATAGCGGCAATGACCTCTGTGATTTCTGCAAT 1560
DB 1501 GAAGACTGGAAGATAAACCTGTTTATAGCGGCAATGACCTCTGTGATTTCTGCAAT 1560
QY 1561 GATCTGCTCAATTTCTCCCGAATTTCAAGACAACTTGGAAAAGCCCTGACATC 1620
DB 1561 GATCTGCTCAATTTCTCCCGAATTTCAAGACAACTTGGAAAAGCCCTGACATC 1620
QY 1621 CTCATGCTGAGGTTCTGCGGCAATTTGGAACCTGGTGAAGGCTGATGATGCAAC 1680
DB 1621 CTCATGCTGAGGTTCTGCGGCAATTTGGAACCTGGTGAAGGCTGATGATGCAAC 1680
QY 1681 CTGAGGAGCTGTACAGAGAGAAAAGTCTACTGCCAAGATGATTCCTCAAGTCTG 1740
DB 1681 CTGAGGAGCTGTACAGAGAGAAAAGTCTACTGCCAAGATGATTCCTCAAGTCTG 1740
QY 1741 TGTCCCTGTGTCTGAAAGTTTGAATTAATCAACGAAGTCTGATCTGATGAAATTC 1800
DB 1741 TGTCCCTGTGTCTGAAAGTTTGAATTAATCAACGAAGTCTGATCTGATGAAATTC 1800
QY 1801 AACAGAAGTTTCAGAGAGAACCAACCACTGATTGAGAGTGGGGATATGACAGAG 1860
DB 1801 AACAGAAGTTTCAGAGAGAACCAACCACTGATTGAGAGTGGGGATATGACAGAG 1860
QY 1861 GAAGATTTTACTGTGTTGTGACAGCGTTCTTTGAAAACGTGACATGCAAGACCTG 1920
DB 1861 GAAGATTTTACTGTGTTGTGACAGCGTTCTTTGAAAACGTGACATGCAAGACCTG 1920
QY 1921 GAAGATTTGCTGACAACTCTTTCTTCTGCTCTGACTGTTTCACTTCAAGCAAGTCT 1980
DB 1921 GAAGATTTGCTGACAACTCTTTCTTCTGCTCTGACTGTTTCACTTCAAGCAAGTCT 1980
QY 1981 CACTCCCGACAGCGCTGCTCTGGAACAAATATCTGAGAGCTGTGGCCGAAGAG 2040
DB 1981 CACTCCCGACAGCGCTGCTCTGGAACAAATATCTGAGAGCTGTGGCCGAAGAG 2040
QY 2041 ACTCGTCATAGTTTGAAGAACAGATCAATATCACTGTCGAACAGGCTCCAGCGCTT 2100
DB 2041 ACTCGTCATAGTTTGAAGAACAGATCAATATCACTGTCGAACAGGCTCCAGCGCTT 2100
QY 2101 CTGAGGACCTTACAGAACAGATGAGGCTCATGAGGACCTGCTGCAATGCAAGAG 2160
DB 2101 CTGAGGACCTTACAGAACAGATGAGGCTCATGAGGACCTGCTGCAATGCAAGAG 2160
QY 2161 GCGCCCTTCTGCTGACCTTACTGATGATGAGGCTGAGACCTGCAAGATCTCAAGT 2220
DB 2161 GCGCCCTTCTGCTGACCTTACTGATGATGAGGCTGAGACCTGCAAGATCTCAAGT 2220
QY 2221 GTGGCTGCTGTGGGGGATCTGTGACGCGTGGCAATGGAATTGGCTCAACAGAG 2280
DB 2221 GTGGCTGCTGTGGGGGATCTGTGACGCGTGGCAATGGAATTGGCTCAACAGAG 2280
QY 2281 CTCGCCGATGTCAACAAGATGCGGGACTGTCAATACAGTGCAGAGGGGAGCGCTCC 2340
DB 2281 CTCGCCGATGTCAACAAGATGCGGGACTGTCAATACAGTGCAGAGGGGAGCGCTCC 2340
QY 2341 CTGGAAGATGTGACCACTTAATCTTCTGCGGAGTTTAAAGAAACCTCAAGGC 2400
DB 2341 CTGGAAGATGTGACCACTTAATCTTCTGCGGAGTTTAAAGAAACCTCAAGGC 2400
QY 2401 TAGCGCTGGGACGCGGTGATGCAATGACAGATGCAATTCCTCAATCAAGCTGTTCC 2460
DB 2401 TAGCGCTGGGACGCGGTGATGCAATGACAGATGCAATTCCTCAATCAAGCTGTTCC 2460
QY 2461 GGAGCAAGGCTGAGGATCTTATAGCAAGTCAAACTCTGATGCAAGATGAAGAAT 2520
DB 2461 GGAGCAAGGCTGAGGATCTTATAGCAAGTCAAACTCTGATGCAAGATGAAGAAT 2520
QY 2521 GATCATAGATTAATTTCCATGAAGCTGGAAGTCAATCAAGTCTGATCGAGGACG 2580
DB 2521 GATCATAGATTAATTTCCATGAAGCTGGAAGTCAATCAAGTCTGATCGAGGACG 2580
QY 2581 GATTATGTGACTACTGACAGATGGAATCTGTATTCGACAGCAACTTGTGACAT 2640

DB 2581 GATTATGTGACTACTGACAGATGGAATCTGTATTCGACAGCAACTTGTGACAT 2640
QY 2641 CTCGCCAATGCTTGGACCTCTTGCATAGAGAGTGGCCAGAGTCTGTGCAACCTG 2700
DB 2641 CTCGCCAATGCTTGGACCTCTTGCATAGAGAGTGGCCAGAGTCTGTGCAACCTG 2700
QY 2701 GACTTCTTAACCCCACTATCAATGCGGAGGTTTCTGGGAAACCCAGCAAGTCCCA 2760
DB 2701 GACTTCTTAACCCCACTATCAATGCGGAGGTTTCTGGGAAACCCAGCAAGTCCCA 2760
QY 2761 GTGACAGAGCCAGCGTTTGTGTAACTGCGTTCTGACCTTGCGGAGAACTTCCAGAG 2820
DB 2761 GTGACAGAGCCAGCGTTTGTGTAACTGCGTTCTGACCTTGCGGAGAACTTCCAGAG 2820
QY 2821 CTAGCCAGGCTGAGGCTTTCACGCGAGCTTACCGGACAGATGCGCGAGCTGTGGG 2880
DB 2821 CTAGCCAGGCTGAGGCTTTCACGCGAGCTTACCGGAGAGATGCGCGAGCTGTGGG 2880
QY 2881 TCAGGCGGCTATGACAGGAGAGGACTTCTGTGTGCTGACGCTTCTTCCAGAAC 2940
DB 2881 TCAGGCGGCTATGACAGGAGAGGACTTCTGTGTGCTGACGCTTCTTCCAGAAC 2940
QY 2941 ATTCAGCTCCCTGTCTGTGCGGATGCGGCTCCAGATAGTCTTCTTTCGCCAGCTG 3000
DB 2941 ATTCAGCTCCCTGTCTGTGCGGATGCGGCTCCAGATAGTCTTCTTTCGCCAGCTG 3000
QY 3001 ATTCAGCCAAATCAGAAATTCATCTCCAGCTGCGGACAGCTTTTGGACCAATAGCTT 3060
DB 3001 ATTCAGCCAAATCAGAAATTCATCTCCAGCTGCGGACAGCTTTTGGACCAATAGCTT 3060
QY 3061 GAACCACTTGAAGCAAAACAGAGACCTGGAACCTGAGAGAGATGACCATCACTGT 3120
DB 3061 GAACCACTTGAAGCAAAACAGAGACCTGGAACCTGAGAGAGATGACCATCACTGT 3120
QY 3121 CCCACTCAGATGAGCCCTTCTGAGAACCCCTCGGAATAGTAACTTACACCTCCATC 3180
DB 3121 CCCACTCAGATGAGCCCTTCTGAGAACCCCTCGGAATAGTAACTTACACCTCCATC 3180
QY 3181 AAGCCAGCATTTGAGAACTGGGGAGTGAATCTCTGTGTACAGAGTGAAGGCTTCCAT 3240
DB 3181 AAGCCAGCATTTGAGAACTGGGGAGTGAATCTCTGTGTACAGAGTGAAGGCTTCCAT 3240
QY 3241 AGTGTTCGAACCTCTGTCCAGAGCTCCGACAGAGCATCAATAGTGTGGCGGCTG 3300
DB 3241 AGTGTTCGAACCTCTGTCCAGAGCTCCGACAGAGCATCAATAGTGTGGCGGCTG 3300
QY 3301 GGTGACTCTGTGACAGAGTGGGAGCTGACCAAACTCCAGTACCTTACCCACA 3360
DB 3301 GGTGACTCTGTGACAGAGTGGGAGCTTCCACCAAACTCCAGTACCTTACCCACA 3360
QY 3361 TCTTGGAGGGGACTCTTGTGAGCATTTGAGAGGGGATGGAACTTGGAGCTCACACCA 3420
DB 3361 TCTTGGAGGGGACTCTTGTGAGCATTTGAGAGGGGATGGAACTTGGAGCTCACACCA 3420
QY 3421 CTGCCCAATTTGTGAAGAGTTTCAACCTTACTCTTGGCTTCTTACAGACCTGG 3480
DB 3421 CTGCCCAATTTGTGAAGAGTTTCAACCTTACTCTTGGCTTCTTACAGACCTGG 3480
QY 3481 GAGGGGACAGAGATTAATGTGAGGAGGAGGAGGAGGAGCTGAGGACATGCGAGCC 3540
DB 3481 GAGGGGACAGAGATTAATGTGAGGAGGAGGAGGAGGAGGAGCTGAGGACATGCGAGCC 3540
QY 3541 CAGGCTGTGGACCTGTGAGCGGAATGAAGAACGCGCCGACATCACTTGAAGAAAG 3600
DB 3541 CAGGCTGTGGACCTGTGAGCGGAATGAAGAAAGCAGCCCGGACATCACTTGAAGAAAG 3600
QY 3601 TGAAGCTGTGACACTTCTTCAATGGGGTCAACGATTTGTGTACTTACTGTGAATTCG 3660
DB 3601 TGAAGCTGTGACACTTCTTCAATGGGGTCAACGATTTGTGTACTTACTGTGAATTCG 3660
QY 3661 GAGGCCCATTTGGGACGGAATATGTTGAGCATCTCAACAGGCGCTGACATCTCTCT 3720

QY 301 CAGGTGTCAGGAGATGATGACAGTCTTTACAGATCATCATATTTCAAGTCTTCT 360
DB 301 CAGGTGTCAGGAGATGATGACAGTCTTTACAGATCATCATATTTCAAGTCTTCT 360
QY 361 GTTTCAGATGCTGTGTGTCACACTGGAAGAGAGTCTATACCCACGATGCTGTAAGAC 420
DB 361 GTTTCAGATGCTGTGTGTCACACTGGAAGAGAGTCTATACCCACGATGCTGTAAGAC 420
QY 421 TTGTGATTCAGGCTCAGAGATGCTGTGAGAAACATGAAAGAACTTGCACCTTGA 480
DB 421 TTGTGATTCAGGCTCAGAGATGCTGTGAGAAACATGAAAGAACTTGCACCTTGA 480
QY 481 CAATTTGATGAGAGTCTCATCATATGTGTCTTCAATATGACAGCAGTGTATCTGTG 540
DB 481 CAATTTGATGAGAGTCTCATCATATGTGTCTTCAATATGACAGCAGTGTATCTGTG 540
QY 541 CCTCTGCTCAACAGAAATGAGGCTTGGCGGCGGCGCTGATGAGTGGGGGTGCTG 600
DB 541 CCTCTGCTCAACAGAAATGAGGCTTGGCGGCGGCGCTGATGAGTGGGGGTGCTG 600
QY 601 GACTTACTGACAGAGAGGTTCCCAAGCATTTGTAACTGTGTGACCTTCTGTAGGTT 660
DB 601 GACTTACTGACAGAGAGGTTCCCAAGCATTTGTAACTGTGTGACCTTCTGTAGGTT 660
QY 661 GCAGAGGCTCTGCTGATGATCAAGGCACTTGGCTCAGCCCTGACAGAGCCCTGTAAT 720
DB 661 GCAGAGGCTCTGCTGATGATCAAGGCACTTGGCTCAGCCCTGACAGAGCCCTGTAAT 720
QY 721 TGCTCAGAGAGAGACCAACCCGCTGCGCCAGAGTGTGATGCAAGTGTCTTATCAGAAAGC 780
DB 721 TGCTCAGAGAGAGACCAACCCGCTGCGCCAGAGTGTGATGCAAGTGTCTTATCAGAAAGC 780
QY 781 TGGAAACAGCTCTCTGAGCTTCCAGAGGATACAGTGAAGAGAGTCTTCAACCGTGTTC 840
DB 781 TGGAAACAGCTCTCTGAGCTTCCAGAGGATACAGTGAAGAGAGTCTTCAACCGTGTTC 840
QY 841 CAGCCTTTCTTCTATGAGACCAACCCCACTCTCACTCGAGAGAGCCCGGACTCCAGAT 900
DB 841 CAGCCTTTCTTCTATGAGACCAACCCCACTCTCACTCGAGAGAGCCCGGACTCCAGAT 900
QY 901 TCTTACCAAGCTGAGCTGTGATCTCTGAGATAGAGTGAATGAGAGCAGAGAGAGAAAT 960
DB 901 TCTTACCAAGCTGAGCTGTGATCTCTGAGATAGAGTGAATGAGAGCAGAGAGAGAAAT 960
QY 961 GAGCCATGAGTGTAAACACGAGAGGCAATGAAATGTCTCTCAGAGAGCCCTTAT 1020
DB 961 GAGCCATGAGTGTGTAAACACGAGAGGCAATGAAATGTCTCTCAGAGAGCCCTTAT 1020
QY 1021 CTGTTACAGCTACAGAAACAGCAATCTACTGACAGACTGAGAAACCCCAAGCAAGCTT 1080
DB 1021 CTGTTACAGCTACAGAAACAGCAATCTACTGACAGACTGAGAAACCCCAAGCAAGCTT 1080
QY 1081 GAGGTAAAGAAAGAGAGGAAATCAGATGTCTGACAAAGACCCCTCCGATACGGTTCC 1140
DB 1081 GAGGTAAAGAAAGAGAGGAAATCAGATGTCTGACAAAGACCCCTCCGATACGGTTCC 1140
QY 1141 ACCTCAGTTCAATGAGTGAAGCCGAGCTGACATCAACGTAAATGAGCCCTGAGTGTCT 1200
DB 1141 ACCTCAGTTCAATGAGTGAAGCCGAGCTGACATCAACGTAAATGAGCCCTGAGTGTCT 1200
QY 1201 CTGACGAGAGGAGTAAGGAGCCGAGTCCACACCTGAGAAAGTCTTGTGACGTCTGAC 1260
DB 1201 CTGACGAGAGGAGTAAGGAGCCGAGTCCACACCTGAGAAAGTCTTGTGACGTCTGAC 1260
QY 1261 TACCGAGGCTGTCTGTGAGAGCTGTGAGAGATGAGAAACATCGGACCGTTACACCTG 1320
DB 1261 TACCGAGGCTGTCTGTGAGAGCTGTGAGAGATGAGAAACATCGGACCGTTACACCTG 1320
QY 1321 GCGAATCTCTCGGGAATTCACCTTCTCTGAGAGGCTTCTCTGTGCACTGAGAA 1380
DB 1321 GCGAATCTCTCGGGAATTCACCTTCTCTGAGAGGCTTCTCTGTGCACTGAGAA 1380

QY 1381 GAAACAGTCTTAATGCTTCTTAAACAGAGCTGTGAGAGAGCCGAGTGAATCTA 1440
DB 1381 GAAACAGTCTTAATGCTTCTTAAACAGAGCTGTGAGAGAGCCGAGTGAATCTA 1440
QY 1441 CCTGTCCAGGCCAGAGAGCTGTGAGACCTGATGAAGATGACACAGAGATACCTTTCAG 1500
DB 1441 CCTGTCCAGGCCAGAGAGCTGTGAGACCTGATGAAGATGACACAGAGATACCTTTCAG 1500
QY 1501 GAAGACTGGAAGATTAATTAACCCGTTTATGAGGGGCAATGACCTGCTGATTTCTGCAAT 1560
DB 1501 GAAGACTGGAAGATTAATTAACCCGTTTATGAGGGGCAATGACCTGCTGATTTCTGCAAT 1560
QY 1561 GATCTGTCCATATTTCTCCCAAGCTTCAACAGAACTTGAAGAGCCCTGAGATC 1620
DB 1561 GATCTGTCCATATTTCTCCCAAGCTTCAACAGAACTTGAAGAGCCCTGAGATC 1620
QY 1621 CTCATGCTGAGGTTCTCGGAGATTTGTGAACCTGTGTGAGCTGTGAGATGTCAC 1680
DB 1621 CTCATGCTGAGGTTCTCGGAGATTTGTGAACCTGTGTGAGCTGTGAGATGTCAC 1680
QY 1681 CTGAGGAGCTGTACAGAGAGAAAGATCTATGCCCCAAGATGATCTCAGAGTCTG 1740
DB 1681 CTGAGGAGCTGTACAGAGAGAAAGATCTATGCCCCAAGATGATCTCAGAGTCTG 1740
QY 1741 TGTCCCTGTGCTCTGAAGTTTGAATGATTAACCTCAAGAACTTGCATCTCAATG 1800
DB 1741 TGTCCCTGTGCTCTGAAGTTTGAATGATTAACCTCAAGAACTTGCATCTCAATG 1800
QY 1801 AACAGAGGTTTCAAGAGAGAAAGCAACCACTGATTTGAGAGTGGGATATGACAAAG 1860
DB 1801 AACAGAGGTTTCAAGAGAGAAAGCAACCACTGATTTGAGAGTGGGATATGACAAAG 1860
QY 1861 GAAGATTTTACTGTGTGTGTGACAGCCGTTTGTGAAGAGTGAACATGCAAGACCTG 1920
DB 1861 GAAGATTTTACTGTGTGTGTGACAGCCGTTTGTGAAGAGTGAACATGCAAGACCTG 1920
QY 1921 GAAGATTTTACTGTGTGTGTGACAGCCGTTTGTGAAGAGTGAACATGCAAGACCTG 1980
DB 1921 GAAGATTTTACTGTGTGTGTGACAGCCGTTTGTGAAGAGTGAACATGCAAGACCTG 1980
QY 1981 CACTCCGAGAGAGAGAGTGTCTCTGAGAAACATATGCTGTGAGAGCTGTGAGAGAG 2040
DB 1981 CACTCCGAGAGAGAGAGTGTCTCTGAGAAACATATGCTGTGAGAGCTGTGAGAGAG 2040
QY 2041 ACTGTCTAATGATTTGAAACAAAGATCATATATCATATGTCGAACAGAGTCCAGCCGTT 2100
DB 2041 ACTGTCTAATGATTTGAAACAAAGATCATATATCATATGTCGAACAGAGTCCAGCCGTT 2100
QY 2101 CTGAGAGACCTACAGAAACAGAGAGTCAAGGAGTCAAGGAGCTGTGAGTCAAGAGAG 2160
DB 2101 CTGAGAGACCTACAGAAACAGAGAGTCAAGGAGTCAAGGAGCTGTGAGTCAAGAGAG 2160
QY 2161 GCCCTTTGCTTGTGACACCTTACCTGATGATGAGCTGTGAGAGCTGTGAGAGAG 2220
DB 2161 GCCCTTTGCTTGTGACACCTTACCTGATGATGAGCTGTGAGAGCTGTGAGAGAG 2220
QY 2221 GTGGTGTCTGTGGGAGATTTCTGTACCGCTGTGAGATGAAATGTGCTCAACAGAGAG 2280
DB 2221 GTGGTGTCTGTGGGAGATTTCTGTACCGCTGTGAGATGAAATGTGCTCAACAGAGAG 2280
QY 2281 CTGCGGAGTGTACACAGATGATGAGGAGCTGTGATGAGTCAAGTCAAGAGAGAGGCTG 2340
DB 2281 CTGCGGAGTGTACACAGATGATGAGGAGCTGTGATGAGTCAAGTCAAGAGAGAGGCTG 2340
QY 2341 CTGGAATATGTGACCACTTACTTAATCTCTGAGAGATTTTAAACAGAACTTCAAGAG 2400
DB 2341 CTGGAATATGTGACCACTTACTTAATCTCTGAGAGATTTTAAACAGAACTTCAAGAG 2400
QY 2401 TACGCGGTGGGAGAGGATGATGCAATGACAGAAATGCAATGCAATGCAATGCAATG 2460
DB 2401 TACGCGGTGGGAGAGGATGATGCAATGACAGAAATGCAATGCAATGCAATGCAATG 2460
QY 2461 GAGAGAAAGGCTGAGAGATCTTATGAGCCAAAGTCCAACTGTGATGAGAAAGATGAAAGAT 2520

Db 2461 GGAGCAAGGCTGAGATCTTATGAGCCAACTCGATGAGAGATGAAGAT 2520
 Qy 2521 GATCATAGATAAATTTCCATGAAAGACTGGAAGTCTACAGTGTCTGATCGAGGAC 2580
 Db 2521 GATCATAGATAAATTTCCATGAAAGACTGGAAGTCTACAGTGTCTGATCGAGGAC 2580
 Qy 2581 GATTTATGACTACTGACAGATTCGATCTGTATCTGAGCAACTTGTGTACCAT 2640
 Db 2581 GATTTATGACTACTGACAGATTCGATCTGTATCTGAGCAACTTGTGTACCAT 2640
 Qy 2641 CTCGCAATGCTTTGAGCGTCTGCAATAGAGAGGTGCGAAGTCTGATCAACTGTG 2700
 Db 2641 CTCGCAATGCTTTGAGCGTCTGCAATAGAGAGGTGCGAAGTCTGATCAACTGTG 2700
 Qy 2701 GACTTCTGAAACCCCACTATCATGCGGAGGTGTCTGTGGAAACCCAGACAGTCCCA 2760
 Db 2701 GACTTCTGAAACCCCACTATCATGCGGAGGTGTCTGTGGAAACCCAGACAGTCCCA 2760
 Qy 2761 GTGAGAGAGGCGAGCGTTTGTGTAACTGGGTTGACCCCTGGGGAGAACTCCCAAG 2820
 Db 2761 GTGAGAGAGGCGAGCGTTTGTGTAACTGGGTTGACCCCTGGGGAGAACTCCCAAG 2820
 Qy 2821 CTAGCCAGGCTGAGGCTTCAAGCCGAGCTACCGGAGCAGATGCGGAGCTGTGGG 2880
 Db 2821 CTAGCCAGGCTGAGGCTTCAAGCCGAGCTACCGGAGCAGATGCGGAGCTGTGGG 2880
 Qy 2881 TCAGGCGGTATGACAGCAGAGAGATCTTCTGTGTGTCTGACGCTTCTTCCAGAC 2940
 Db 2881 TCAGGCGGTATGACAGCAGAGAGATCTTCTGTGTGTCTGACGCTTCTTCCAGAC 2940
 Qy 2941 ATCAGCTCCCTGCTGCGGAGATGGGCTCCAGATAGTCTTCTTGGCCGAGATGC 3000
 Db 2941 ATCAGCTCCCTGCTGCGGAGATGGGCTCCAGATAGTCTTCTTGGCCGAGATGC 3000
 Qy 3001 ATCCACCCCAATGAGAAATTCATCTCCAGCTGGCCAGAGCCCTTGGACCAATATGCTT 3060
 Db 3001 ATCCACCCCAATGAGAAATTCATCTCCAGCTGGCCAGAGCCCTTGGACCAATATGCTT 3060
 Qy 3061 GAACCACTTGGAGCAAAAGAGAGATCTTCTGTGTGTCTGAGAGAGATGCCATATCCTGT 3120
 Db 3061 GAACCACTTGGAGCAAAAGAGAGATCTTCTGTGTGTCTGAGAGAGATGCCATATCCTGT 3120
 Qy 3121 CCCACTGAGATGAGCCCTTCTGAGAACTCTGGAATGTATCTACAGTATCCCATC 3180
 Db 3121 CCCACTGAGATGAGCCCTTCTGAGAACTCTGGAATGTATCTACAGTATCCCATC 3180
 Qy 3181 AAGCCAGCACTTGAAGATGAGGAGCTGATCTTCTGTGTATCAGAGTGAAGGCTTCCAAT 3240
 Db 3181 AAGCCAGCACTTGAAGATGAGGAGCTGATCTTCTGTGTATCAGAGTGAAGGCTTCCAAT 3240
 Qy 3241 AGTGTTCGAACCTGTGTCCACAGCTCCGACAGCAGACATCAAGTGTGGCCGCTG 3300
 Db 3241 AGTGTTCGAACCTGTGTCCACAGCTCCGACAGCAGACATCAAGTGTGGCCGCTG 3300
 Qy 3301 GGTAACCTCTGCTACAGAGATGGGAGCTCCAGCAAACTCCATGACCTTACCCACA 3360
 Db 3301 GGTAACCTCTGCTACAGAGATGGGAGCTCCAGCAAACTCCATGACCTTACCCACA 3360
 Qy 3361 TCTTGAAGGGAATCTTCTGAGCACTTGAAGGAGTGGAACTTGAAGCTCAACACCA 3420
 Db 3361 TCTTGAAGGGAATCTTCTGAGCACTTGAAGGAGTGGAACTTGAAGCTCAACACCA 3420
 Qy 3421 CTGCCCCAATCTTGAAGAGTTCACCTTACTCTTGGCTTCTTACACAGACCTGG 3480
 Db 3421 CTGCCCCAATCTTGAAGAGTTCACCTTACTCTTGGCTTCTTACACAGACCTGG 3480
 Qy 3481 GAGGGGACAGAGACTTAAATGTGGCAGCGGAAGGGCCAGAGCTAAGGACATGCCAGCC 3540
 Db 3481 GAGGGGACAGAGACTTAAATGTGGCAGCGGAAGGGCCAGAGCTAAGGACATGCCAGCC 3540
 Qy 3541 CAGGCTCTGGACCTGTAGAGCAGATGAAGAAACAGCCCGACATCACTGAGAGAAAGAC 3600
 Db 3541 CAGGCTCTGGACCTGTAGAGCAGATGAAGAAACAGCCCGACATCACTGAGAGAAAGAC 3600

Db 3541 CAGGCTCTGGACCTGTAGAGCAGATGAAGAAACAGCCCGACATCACTGAGAGAAAGAC 3600
 Qy 3601 TGGAGCTGTGACACTCTTCAATTTGGGGTCAACGACTTGTGTATTAAGTGAATCCG 3660
 Db 3601 TGGAGCTGTGACACTCTTCAATTTGGGGTCAACGACTTGTGTATTAAGTGAATCCG 3660
 Qy 3661 GAGGCCCACTTGGCCACGGAATATGTTCAGCATCAACAGGCCCTGAGCATCTCTCT 3720
 Db 3661 GAGGCCCACTTGGCCACGGAATATGTTCAGCATCAACAGGCCCTGAGCATCTCTCT 3720
 Qy 3721 GAGAGCTCCCAAGGCTTCTGCAACGTGTGAGAGTCAATGAGTCTGCTTACCTGTAC 3780
 Db 3721 GAGAGCTCCCAAGGCTTCTGCAACGTGTGAGAGTCAATGAGTCTGCTTACCTGTAC 3780
 Qy 3781 CAGGGCCAGGGGGGAAATGTGCATGTGGGAGCTCAGAACTGCACTTGGCTCAGA 3840
 Db 3781 CAGGGCCAGGGGGGAAATGTGCATGTGGGAGCTCAGAACTGCACTTGGCTCAGA 3840
 Qy 3841 CACTCCCAAGCTCCCTGAGAGAGCAAGATGAAAGTGAATGGAACCTCCAGCAT 3900
 Db 3841 CACTCCCAAGCTCCCTGAGAGAGCAAGATGAAAGTGAATGGAACCTCCAGCAT 3900
 Qy 3901 GGCATCTCCAGTTTCTCTTACCTGAGCACTAATACACAGCTGAGAGACTTGTGTGTG 3960
 Db 3901 GGCATCTCCAGTTTCTCTTACCTGAGCACTAATACACAGCTGAGAGACTTGTGTGTG 3960
 Qy 3961 GTGCAAGCTTCTTCCAAACACATCACCCTCACTGAAGAGAGAGGAGACATGACCTC 4020
 Db 3961 GTGCAAGCTTCTTCCAAACACATCACCCTCACTGAAGAGAGAGGAGACATGACCTC 4020
 Qy 4021 ACCCTCTTCCAGAGACTGTTTCACTTCTCAGACCGGAGGATCCGAGATGAGCATC 4080
 Db 4021 ACCCTCTTCCAGAGACTGTTTCACTTCTCAGACCGGAGGATCCGAGATGAGCATC 4080
 Qy 4081 GCACTCTGGAACCAATGCTGGAACCAAGTGGGCGGAGACTTCTCAACCACTTACCC 4140
 Db 4081 GCACTCTGGAACCAATGCTGGAACCAAGTGGGCGGAGACTTCTCAACCACTTACCC 4140
 Qy 4141 CAGAGCCAGCCCAACTCAAGTGGCCCTCTCTGAGAGCCCTTACCTACACCTGGCGG 4200
 Db 4141 CAGAGCCAGCCCAACTCAAGTGGCCCTCTCTGAGAGCCCTTACCTACACCTGGCGG 4200
 Qy 4201 AACAGCCGATGTCTCCAGACAGCTGGAAGAGCCCGAGGCTGCTTACTGAGGCTGC 4260
 Db 4201 AACAGCCGATGTCTCCAGACAGCTGGAAGAGCCCGAGGCTGCTTACTGAGGCTGC 4260
 Qy 4261 CCAAGTGGCAGCGGAGATCGGCTTGTGTGGGATCATCGGAGCAATGTGTGGAGTGC 4320
 Db 4261 CCAAGTGGCAGCGGAGATCGGCTTGTGTGGGATCATCGGAGCAATGTGTGGAGTGC 4320
 Qy 4321 AGGAGAGGTGGCCGAGGGAAGATCTCTCAATAGCTTGGCATCTGTGGCCCTTACG 4377
 Db 4321 AGGAGAGGTGGCCGAGGGAAGATCTCTCAATAGAGCTTGGCATCTGTGGCCCTTACG 4377

RESULT 3
 AAD52626
 ID AAD52626 standard; cDNA; 4607 BP.
 XX AAD52626;
 DT 14-MAY-2003 (first entry)
 DE Human lipid-associated molecule (LIPAM) -1 cDNA.
 XX Human; lipid-associated molecule; LIPAM; cardiovascular disorder; stroke;
 KW arteriovenous fistula; atherosclerosis; hypertension; Raynaud's disease;
 KW aneurysm; congestive heart failure; thrombophlebitis; angina pectoris;
 KW ischemic heart disease; rheumatic heart disease; peptic esophagitis;
 KW gastrointestinal disorder; lipid metabolism disorder; Crohn's disease;
 KW nausea; peptic ulcer; fatty liver; Fabry's disease; Gaucher's disease;
 KW diabetes mellitus; hyperlipidemia; hypercholesterolemia; epilepsy;
 KW autoimmune disorder; inflammatory disorder; neurological disorder; kuru;

QY 781 TGAACAGCCTCCGCGCTCCAGAGGTACGTGAGCAGGAGTCTTACCGTGTTC 840
DB 789 TGAACAGCCTCCGCGCTCCAGAGGTACGTGAGCAGGAGTCTTACCGTGTTC 848
QY 841 CAGCCTTCTTCTATGAGACCACTCTCTCACTCGAGAGACCCCGACTCGAGAT 900
DB 849 CAGCCTTCTTCTATGAGACCACTCTCTCACTCGAGAGACCCCGACTCGAGAT 908
QY 901 TCTACCACTGCGCTGCGCATCTCTGGAATGATGATGAGCCAGCAGAGAGAAAT 960
DB 909 TCTACCACTGCGCTGCGCATCTCTGGAATGATGATGAGCCAGCAGAGAGAAAT 968
QY 961 GAGCACTTGAATGAAACAGCGGAGGCAATGAGTGCCTCTCGAGAGACCCCTAT 1020
DB 969 GAGCACTTGAATGAAACAGCGGAGGCAATGAGTGCCTCTCGAGAGACCCCTAT 1028
QY 1021 CTGTTCACTTACAGAAACAGCACTTACCTGACAGCTGAGAAACCCCAAGCAAGCTT 1080
DB 1029 CTGTTCACTTACAGAAACAGCACTTACCTGACAGCTGAGAAACCCCAAGCAAGCTT 1088
QY 1081 GAGTAAAG 1140
DB 1089 GAGTAAAG 1148
QY 1141 ACCTCACTTATAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1149 ACCTCACTTATAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1208
QY 1201 CTGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1260
DB 1209 CTGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1268
QY 1261 TACCGAGGCTGTCTGAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1269 TACCGAGGCTGTCTGAGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1328
QY 1321 GCGAACACTCTCGGAGATCAACCTCTCTCGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB 1329 GCGAACACTCTCGGAGATCAACCTCTCTCGAGAGAGAGAGAGAGAGAGAGAGAG 1388
QY 1381 GAAACAGCTCTATGCTCTTAAACAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAT 1440
DB 1389 GAAACAGCTCTATGCTCTTAAACAGAGCTGTGAGAGAGAGAGAGAGAGAGAT 1448
QY 1441 CCGTCCAGGAGCAGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1500
DB 1449 CCGTCCAGGAGCAGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1508
QY 1501 GAGGACCTGGAAGATTAATACCTGTTATAGGAGGAGAGAGAGAGAGAGAGAGAT 1560
DB 1509 GAGGACCTGGAAGATTAATACCTGTTATAGGAGGAGAGAGAGAGAGAGAGAT 1568
QY 1561 GATCTGTCTCACTATCTCTCCAGAACTTCAAGAGAGAGAGAGAGAGAGAGAGATC 1620
DB 1569 GATCTGTCTCACTATCTCTCCAGAACTTCAAGAGAGAGAGAGAGAGAGAGATC 1628
QY 1621 CTCCATCTGAGGTTCTCGGAGATTTGTGAACCTGTGAGAGAGAGAGAGAGAT 1680
DB 1629 CTCCATCTGAGGTTCTCGGAGATTTGTGAACCTGTGAGAGAGAGAGAGAT 1688
QY 1681 CTGAGGAGAGCTGTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1740
DB 1689 CTGAGGAGAGCTGTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1748
QY 1741 TGTCTCTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1800
DB 1749 TGTCTCTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1808
QY 1801 AACAGAT 1860
DB 1809 AACAGAT 1868

QY 1861 GAAGATTTTACTGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1920
DB 1869 GAAGATTTTACTGTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1928
QY 1921 GAGAGATGCTGAGACATCTTCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAT 1980
DB 1929 GAGAGATGCTGAGACATCTTCTTCTGCTGAGAGAGAGAGAGAGAGAGAT 1988
QY 1981 CATCTCCAGAT 2040
DB 1989 CATCTCCAGAT 2048
QY 2041 ACTGCTCATAGTTGAAACAGAGATCAATATCATAGTCCAGAGAGAGAGAT 2100
DB 2049 ACTGCTCATAGTTGAAACAGAGATCAATATCATAGTCCAGAGAGAGAT 2108
QY 2101 CTGAT 2160
DB 2109 CTGAT 2168
QY 2161 GCGCCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2220
DB 2169 GCGCCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2228
QY 2221 GTGCTGCTCTGAGAGAGATCTCTGAGAGAGAGAGAGAGAGAGAGAT 2280
DB 2229 GTGCTGCTCTGAGAGAGATCTCTGAGAGAGAGAGAGAGAGAGAT 2288
QY 2281 CTCGCGAT 2340
DB 2289 CTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2348
QY 2341 CTGAT 2400
DB 2349 CTGAT 2408
QY 2401 TACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2460
DB 2409 TACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2468
QY 2461 GAT 2520
DB 2469 GAT 2528
QY 2521 GATCATAGAGATTAATTTCAAGAGAGAGAGAGAGAGAGAGAGAT 2580
DB 2529 GATCATAGAGATTAATTTCAAGAGAGAGAGAGAGAGAGAGAGAT 2588
QY 2581 GATTTATGATCACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2640
DB 2589 GATTTATGATCACTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2648
QY 2641 CTCGCAATGCTTGAAGCTCTGAGAGAGAGAGAGAGAGAGAGAT 2700
DB 2649 CTCGCAATGCTTGAAGCTCTGAGAGAGAGAGAGAGAGAGAT 2708
QY 2701 GACTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2760
DB 2709 GACTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2768
QY 2761 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2820
DB 2769 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2828
QY 2821 CTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2880
DB 2829 CTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2888
QY 2881 TCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2940
DB 2889 TCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2948
QY 2941 ATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 3000

```

Db      ||| 2949 ATCCAGTCCCTGCTGCGGAGATGGGCTCCAGATCGTCTTCTTGGCCCGAGACTGC 3008
Qy      ||| 3001 ATCCACCAATGAGAAATTCATCCAGCTGGCCGAGACCCCTTGGACCAATAGCTT 3060
Db      ||| 3009 ATCCACCAATGAGAAATTCATCCAGCTGGCCGAGACCCCTTGGACCAATAGCTT 3068
Qy      ||| 3061 GAACCACTTGAAGCAAAACAGAGACCTTGAGACCTGAGACAGATGCCATCAGCTT 3120
Db      ||| 3069 GAACCACTTGAAGCAAAACAGAGACCTTGAGACCTGAGACAGATGCCATCAGCTT 3128
Qy      ||| 3121 CCCACTGAGATAGCCCTTCTGAGAACCCCTGGAATAGTAACTAACGTAACCCATC 3180
Db      ||| 3129 CCCACTGAGATAGCCCTTCTGAGAACCCCTGGAATAGTAACTAACGTAACCCATC 3188
Qy      ||| 3181 AAGCCAGCACTTGAAGCTGGGGGAGTGAATTCCTGTGTACAGAGTGAAGGCTTCCAAT 3240
Db      ||| 3189 AAGCCAGCACTTGAAGCTGGGGGAGTGAATTCCTGTGTACAGAGTGAAGGCTTCCAAT 3248
Qy      ||| 3241 AGTGTTCGAACCTCTGTCCACAGCTCCGACAGACATCAAAAGTGTGGCCGCTTG 3300
Db      ||| 3249 AGTGTTCGAACCTCTGTCCACAGCTCCGACAGACATCAAAAGTGTGGCCGCTTG 3308
Qy      ||| 3301 GGTGACTCTTGTACTACAGAGTGGGAGTGCACAAACAACTCCAGTGAATCCACA 3360
Db      ||| 3309 GGTGACTCTTGTACTACAGAGTGGGAGTGCACAAACAACTCCAGTGAATCCACA 3368
Qy      ||| 3361 TCTTGAAGGGGAGTCTCTTGGAGCATTTGAGGGGATGGAACTTGAAGATCAACACCA 3420
Db      ||| 3369 TCTTGAAGGGGAGTCTCTTGGAGCATTTGAGGGGATGGAACTTGAAGATCAACACCA 3428
Qy      ||| 3421 CTGCCCAACATTTCTGAAGAAATCAACCCCTTACCTCTTCTCTTCTTACCAACACCTG 3480
Db      ||| 3429 CTGCCCAACATTTCTGAAGAAATCAACCCCTTACCTCTTCTCTTCTTACCAACACCTG 3488
Qy      ||| 3481 GAGGGGACAGCAGAGACTAAATGTGGACCGGAAGGGGCCAGACTGAGGACATGCCAGCC 3540
Db      ||| 3489 GAGGGGACAGCAGAGACTAAATGTGGACCGGAAGGGGCCAGACTGAGGACATGCCAGCC 3548
Qy      ||| 3541 CAGGCTTGGGACCTGTGAGAGCAATGAAAAACAGCCCGCACTCAACCTTGAAGAAAGAC 3600
Db      ||| 3549 CAGGCTTGGGACCTGTGAGAGCAATGAAAAACAGCCCGCACTCAACCTTGAAGAAAGAC 3608
Qy      ||| 3601 TGGAGCTGCTGACACTCTTCAATTTGGGGTCAAGCACTGTGTATTAAGTGAATCCG 3660
Db      ||| 3609 TGGAGCTGCTGACACTCTTCAATTTGGGGTCAAGCACTGTGTATTAAGTGAATCCG 3668
Qy      ||| 3661 GAGGCCCACTTGGCAACGAAATATGTTACGACATCCACAGGCCCTGACATCTCTCT 3720
Db      ||| 3669 GAGGCCCACTTGGCAACGAAATATGTTACGACATCCACAGGCCCTGACATCTCTCT 3728
Qy      ||| 3721 GAGGAGCTCCCAAGGGCTTCTGCAACGTGTGAGAGTCATGAGCTGCTGACCTGTAC 3780
Db      ||| 3729 GAGGAGCTCCCAAGGGCTTCTGCAACGTGTGAGAGTCATGAGCTGCTGACCTGTAC 3788
Qy      ||| 3781 CAGGGCCAGGGGGGAAATGTGCAAGCTGGGAGCTCAAGAACTGCACTTGGCTTCA 3840
Db      ||| 3789 CAGGGCCAGGGGGGAAATGTGCAAGCTGGGAGCTCAAGAACTGCACTTGGCTTCA 3848
Qy      ||| 3841 CACTGCAAGAGCTCCTGAGAGAAAGAACTGAAGAAAGTGAACCTCCAGCAT 3900
Db      ||| 3849 CACTGCAAGAGCTCCTGAGAGAAAGAACTGAAGAAAGTGAACCTCCAGCAT 3908
Qy      ||| 3901 GGCATCTCCAGTTTCTCTACTAGTGGCAACATACACAGGTGAGAGACTTTGGCGTTGTG 3960
Db      ||| 3909 GGCATCTCCAGTTTCTCTACTAGTGGCAACATACACAGGTGAGAGACTTTGGCGTTGTG 3968
Qy      ||| 3961 GTGCAGAGCTTTCTTCAAAAACACATCAACCCCACTGAAGAGAGAGGACATGACCTC 4020
Db      ||| 3969 GTGCAGAGCTTTCTTCAAAAACACATCAACCCCACTGAAGAGAGAGGACATGACCTC 4028
Qy      ||| 4021 ACCTTCTTCCAGAGAGTGTTCATCTCAGACCGGCGGAGCATCCGAGATGCGCATC 4080

```

```

Db      ||| 4029 ACCTTCTTCCGAGAGCTGTTTCACTTCTCAGACCGGCGGAGCATGCCATGCGCATC 4088
Qy      ||| 4081 GCACTCTGAACAACATGCTGGAACAGTGGGCGCGAAGACTACCTCCAACTTCAACC 4140
Db      ||| 4089 GCACTCTGAACAACATGCTGGAACAGTGGGCGCGAAGACTACCTCCAACTTCAACC 4148
Qy      ||| 4141 CACAGCCGAGCCCAACTCAAGTCCCTCTCTGAGAGGCTTACCTTACACCTTGGCG 4200
Db      ||| 4149 CACAGCCGAGCCCAACTCAAGTCCCTCTCTGAGAGGCTTACCTTACACCTTGGCG 4208
Qy      ||| 4201 AACAGCCGATTTGCTCCCAACAGAGCTGAAGAAAGCCCGAGGTGCTTACTGAGGCTGTC 4260
Db      ||| 4209 AACAGCCGATTTGCTCCCAACAGAGCTGAAGAAAGCCCGAGGTGCTTACTGAGGCTGTC 4268
Qy      ||| 4261 CCAGTGGCAGCGGAGATCGGCTTGTGTGGGATCATGAGGACATGCTGAGAGTGC 4320
Db      ||| 4269 CCAGTGGCAGCGGAGATCGGCTTGTGTGGGATCATGAGGACATGCTGAGAGTGC 4328
Qy      ||| 4321 AGGAGAGTGGCCGAGGAGAAAGATCCTCCAAATGAGCTGCGCATGTGCGCTCTAG 4377
Db      ||| 4329 AGGAGAGTGGCCGAGGAGAAAGATCCTCCAAATGAGCTGCGCATGTGCGCTCTAG 4385

RESULT 4
AAD52634
AAD52634 standard; cDNA; 4424 BP.
AC
AC
AAD52634;
XX
XX
14-MAY-2003 (first entry)
DE
XX
Human lipid-associated molecule (LIPAM)-9 cDNA.
XX
XX
Human; lipid-associated molecule; LIPAM; cardiovascular disorder; stroke;
arteriovenous fistula; atherosclerosis; hypertension; Raynaud's disease;
aneurysm; congestive heart failure; thrombophlebitis; angina pectoris;
ischemic heart disease; rheumatic heart disease; peptic esophagitis;
gastrointestinal disorder; lipid metabolism disorder; Crohn's disease;
nausea; peptic ulcer; fatty liver; Fabry's disease; Gaucher's disease;
diabetes mellitus; hyperlipidaemia; hypercholesterolaemia; epilepsy;
autoimmune disorder; inflammatory disorder; neurological disorder; kuru;
acquired immunodeficiency syndrome; anaemia; Alzheimer's disease; asthma;
dementia; prion disease; Creutzfeldt-Jakob disease; leukaemia; cancer;
adenocarcinoma; lymphoma; melanoma; myeloma; sarcoma; gene therapy;
KM
protein replacement therapy; gene; ss.
XX
OS
Homo sapiens.
XX
XX
FH
Key Location/Qualifiers
FT CDS 9..2918
FT FT /tag= a
FT FT /product= "Human LIPAM-9 protein"
FT FT /tag= b
FT FT /tag= c
FT FT /product= "Human mature LIPAM-9 protein"
FT FT /tag= d
FT FT /tag= e
FT FT /product= "Human mature LIPAM-9 protein"
FT FT /tag= f
FT FT /tag= g
FT FT /product= "Human mature LIPAM-9 protein"
FT FT /tag= h
FT FT /tag= i
FT FT /product= "Human mature LIPAM-9 protein"
FT FT
XX

```

PN WO20029498-A2.
 XX 28-NOV-2002.
 XX 17-MAY-2002; 2002MO-US15688.
 XX 18-MAY-2001; 2001US-292242P.
 PR 25-MAY-2001; 2001US-293726P.
 PR 01-JUN-2001; 2001US-295346P.
 PR 06-JUL-2001; 2001US-303404P.
 PR 24-AUG-2001; 2001US-314754P.
 PR 22-JAN-2002; 2002US-351262P.
 PR 29-MAR-2002; 2002US-368799P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Tang YT, Yue H, Azimzai Y, Baughn MR, Burford N, Reddy R,
 PI Walia NK, Das D, Nguyen DB, Yao MG, Arvizu CS, Lu Y, Gandhi AR,
 PI Griffin JA, Elliott VS, Ramkumar J, Lai PG, Lu DM, Lee EX,
 PI Lee SY, Yue H, Yang J, Triboley CM, Kable AE, Swarnakar A;
 XX MPI; 2003-12079/11.
 DR P-PSDB; AAE34448.
 PT New human lipid-associated molecule (LIPAM) proteins and
 PT polynucleotides, useful for diagnosing, treating or preventing
 PT cardiovascular disorders (e.g. aneurysms), neurological disorders (e.g.
 PT Parkinson's disease) or cancers -
 XX Claim 74; Page 168-169; 171pp; English.
 XX The present invention relates to novel human lipid-associated molecules
 CC (LIPAM) and polynucleotides encoding such proteins. Sequences of the
 CC invention are useful for treating diseases or conditions associated with
 CC decreased expression of functional LIPAM. The antagonist is useful for
 CC treating a disease or condition associated with the overexpression of
 CC functional LIPAM. They are useful for diagnosing, treating or preventing
 CC cardiovascular disorders (e.g. arteriovenous fistula, atherosclerosis,
 CC hypertension, Raynaud's disease, aneurysms, varicose veins, congestive
 CC heart failure, thrombophlebitis, angina pectoris, ischaemic heart disease
 CC or rheumatic heart disease), gastrointestinal disorders (e.g. peptic
 CC oesophagitis, nausea, peptic ulcer or Crohn's disease), lipid metabolism
 CC disorders (e.g. fatty liver, Fabry's disease, Gaucher's disease, diabetes
 CC mellitus, hyperlipidaemia, hypercholesterolaemia), autoimmune disorders
 CC or inflammatory disorders (e.g. acquired immunodeficiency syndrome,
 CC anaemia, asthma or Crohn's disease), neurological disorders (e.g. stroke,
 CC epilepsy, dementia, Alzheimer's disease, or prion diseases such as kuru
 CC or Creutzfeldt-Jakob disease) or cancers (e.g. adenocarcinoma, leukaemia,
 CC lymphoma, melanoma, myeloma or sarcoma). They are also used in gene
 CC therapy and protein replacement therapy. The present sequence is human
 CC LIPAM-9 cDNA.
 XX
 SQ Sequence 4424 BP; 1119 A; 1231 C; 1143 G; 931 T; 0 other;
 Query Match 91.0%; Score 3983.4; DB 25; Length 4424;
 Best Local Similarity 95.6%; Pred. No. 0;
 Matches 4184; Conservative 0; Mismatches 6; Indels 187; Gaps 1;

QY 241 ATTCTCCAGACCCAGAGAGGGGCGATCTGAGAAACAAGACTGAGTAAAGGCCACAG 300
 DB 249 ATTCTCCAGACCCAGAGAGGGGCGATCTGAGAAACAAGACTGAGTAAAGGCCACAG 308
 QY 301 CAGGTGTCATGGAGATGACAGTCTTTTTCAGATCATCATGATATTTCACTCTTCT 350
 DB 309 CAGGTGTCATGGAGATGACAGTCTTTTTCAGATCATCATGATATTTCACTCTTCT 358
 QY 361 GTTCCATATGCTGTGTGCGCACTGGAAGAGATATCCCAAGATGTGTGAAGC 420
 DB 369 GTTCCATATGCTGTGTGCGCACTGGAAGAGATATCCCAAGATGTGTGAAGC 428
 QY 421 TTGTGATTCAGGCTCAAGAACTGTGAGAAACATGAAAGAACTGTGAACTTT 480
 DB 429 TTGTGATTCAGGCTCAAGAACTGTGAGAAACATGAAAGAACTGTGAACTTT 488
 QY 481 CAATTTGACTGGAAGCTCATATGTTCTTTCAGTAAATGCAAGCCAGTAACTGTGC 540
 DB 489 CAATTTGACTGGAAGCTCATATGTTCTTTCAGTAAATGCAAGCCAGTAACTGTGC 548
 QY 541 CCTCTGCTCAGACAGATGAGCTTGTGCGGCGGCGGCTGATGATGAGGCGTGTG 600
 DB 549 CCTCTGCTCAGACAGATGAGCTTGTGCGGCGGCGGCTGATGATGAGGCGTGTG 608
 QY 601 GACTACCTGAGAGAGAGTCCCAAGAGCTTTGTAACCTGTGTGAGCTTGTGAGTT 660
 DB 609 GACTACCTGAGAGAGAGTCCCAAGAGCTTTGTAACCTGTGTGAGCTTGTGAGTT 668
 QY 661 GCAGAGGTCCTGTCAGATATCAAGGCACTTGTGCTCAGCCCTGACAGAGCCCTGTAAT 720
 DB 669 GCAGAGGTCCTGTCAGATATCAAGGCACTTGTGCTCAGCCCTGACAGAGCCCTGTAAT 728
 QY 721 TGTCTAGAGAGACCAACCGGCTGTGCGCAAGGTGTATGATGATGATGATGAGAAAGC 780
 DB 729 TGTCTAGAGAGACCAACCGGCTGTGCGCAAGGTGTATGATGATGATGATGAGAAAGC 788
 QY 781 TGAACAGGCTCTGAGCTTCAGACAGATACAGTAAAGAGAGTCTTCAACCGGTGTTTC 840
 DB 789 TGAACAGGCTCTGAGCTTCAGACAGATACAGTAAAGAGAGTCTTCAACCGGTGTTTC 848
 QY 841 CAGCCTTTCTTATGAGACCAACCCATCTTCACTGTGAGAGACCCCGACTCCAGAT 900
 DB 849 CAGCCTTTCTTATGAGACCAACCCATCTTCACTGTGAGAGACCCCGACTCCAGAT 908
 QY 901 TCTACACGCTGTGCTGTGATCTTGTGATATGATGATGATGATGATGATGATGATGAT 960
 DB 909 TCTACACGCTGTGCTGTGATCTTGTGATATGATGATGATGATGATGATGATGATGAT 968
 QY 961 GAGCCATGAGTAAACACCGGAGGCAATGAAATGTCCTTCAAGAGAGACCCCTAT 1020
 DB 969 GAGCCATGAGTAAACACCGGAGGCAATGAAATGTCCTTCAAGAGAGACCCCTAT 1028
 QY 1021 CTGTTCAGCTACAGAAACAGCACTACCTGACAGACTGACAGAAACCCCAAGCAAGCTT 1080
 DB 1029 CTGTTCAGCTACAGAAACAGCACTACCTGACAGACTGACAGAAACCCCAAGCAAGCTT 1088
 QY 1081 GAGGTAAAGAGAGAGAGGAAATCAATGTTCTTGAACAAAGACCCCTTCCATAGGTTCCC 1140
 DB 1089 GAGGTAAAGAGAGAGAGGAAATCAATGTTCTTGAACAAAGACCCCTTCCATAGGTTCCC 1148
 QY 1141 ACCTCAGTTCATAGGCTGAAAGCGGCTGACATCAAGTAAATTTGAGACCCCTGAGTACTT 1200
 DB 1149 ACCTCAGTTCATAGGCTGAAAGCGGCTGACATCAAGTAAATTTGAGACCCCTGAGTACTT 1208
 QY 1201 CTCACGAGAGCAATGGGCGCGGCTCACACTGTGGAACGTTTGTGACGTTGACTCAG 1260
 DB 1209 CTCACGAGAGCAATGGGCGCGGCTCACACTGTGGAACGTTTGTGACGTTGACTCAG 1268
 QY 1261 TACCGAGGCTGTCTGTGAGCGTGTGCGGAGATGAGAACTGTGACCGTTTACACCTTG 1320
 DB 1269 TACCGAGGCTGTCTGTGAGCGTGTGCGGAGATGAGAACTGTGACCGTTTACACCTTG 1328

QY 1321 GCGAATCTCTCGGGAATTCACCTTCCCTGAAGGGCTTCTGTGGCACTGGGAAA 1380
DB 1329 GCGAATCTCTCGGGAATTCACCTTCCCTGAAGGGCTTCTGTGGCACTGGGAAA 1388
QY 1381 GAAACCACTCTTAATGCTTTTAAACAGAGCTGTGGCAGAGAGCCGAGCTGAGATCTA 1440
DB 1389 GAAACCACTCTTAATGCTTTTAAACAGAGCTGTGGCAGAGAGCCGAGCTGAGATCTA 1448
QY 1441 CCGTCCAGGCGCAGAGGCTGTGGACCTGATGAAGATGACAAGAGATACCTTTCAG 1500
DB 1449 CCGTCCAGGCGCAGAGGCTGTGGACCTGATGAAGATGACAAGAGATACCTTTCAG 1508
QY 1501 GAAGACTGGAAGATTAATACCTGTTTATAGCGCGCAATGACCTGTGTATTTCTGCAAT 1560
DB 1509 GAAGACTGGAAGATTAATACCTGTTTATAGCGCGCAATGACCTGTGTATTTCTGCAAT 1568
QY 1561 GATCTGTGCACTATTCTCCCAAACTTCAACAGACAACTTGGAAAAGCCCTGGACATC 1620
DB 1569 GATCTGTGCACTATTCTCCCAAACTTCAACAGACAACTTGGAAAAGCCCTGGACATC 1628
QY 1621 CTCATGCTGAGGCTCTCGGGATTTGTGACCTGTGTAGCGGCTTGGATGCTCAAC 1680
DB 1629 CTCATGCTGAGGCTCTCGGGATTTGTGACCTGTGTAGCGGCTTGGATGCTCAAC 1688
QY 1681 CTGAGGAGCTGTACAGAGAGAAAAAGTCTACTGCCCAGATGATCTCAGSTCTG 1740
DB 1689 CTGAGGAGCTGTACAGAGAGAAAAAGTCTACTGCCCAGATGATCTCAGSTCTG 1748
QY 1741 TGTCCCTGTCTCTGAAGTTTGAATGATTACTCAACAGAACTTGCCTCATGCAATTC 1800
DB 1749 TGTCCCTGTCTCTGAAGTTTGAATGATTACTCAACAGAACTTGCCTCATGCAATTC 1808
QY 1801 AACAAGAGTTTGAAGAGAGAACCCCACTGATGAGTGGGGAATGTGACAAAG 1860
DB 1809 AACAAGAGTTTGAAGAGAGAACCCCACTGATGAGTGGGGAATGTGACAAAG 1868
QY 1861 GAAGATTTTACTGTGTGTGACAGCGTCTTTGAAAACGTGACATGCCAAAGACTCG 1920
DB 1869 GAAGATTTTACTGTGTGTGACAGCGTCTTTGAAAACGTGACATGCCAAAGACTCG 1928
QY 1921 GAAGATTTGCTGACAACTCTTTCTTGTGCTCTGACTGTTTCACTTCAAGCAGCACT 1980
DB 1929 GAAGATTTGCTGACAACTCTTTCTTGTGCTCTGACTGTTTCACTTCAAGCAGCACT 1988
QY 1981 CATCTCCGAGACCACTGTCTCTGGAACAATATGTGTGAGGCTGTGGCCAGAAAGC 2040
DB 1989 CATCTCCGAGACCACTGTCTCTGGAACAATATGTGTGAGGCTGTGGCCAGAAAGC 2048
QY 2041 ACTGCTCATTAAGTTTGAAGAACAGATCATATCATGATGCGAAGCAGGTCAGCGTTT 2100
DB 2049 ACTGCTCATTAAGTTTGAAGAACAGATCATATCATGATGCGAAGCAGGTCAGCGTTT 2108
QY 2101 CTGAGGACTTACAGAAACAGCATGAGGCTCATGAGGACTGTGGCTGCCATGAGGAGAGA 2160
DB 2109 CTGAGGACTTACAGAAACAGCATGAGGCTCATGAGGACTGTGGCTGCCATGAGGAGAGA 2168
QY 2161 GCGCTTCTGTGCTTGGACCTTACCTGATGATGCTGTGAGACCTGACATCAAGTT 2220
DB 2169 GCGCTTCTGTGCTTGGACCTTACCTGATGATGCTGTGAGACCTGACATCAAGTT 2228
QY 2221 GTGGCTGCTGTGGGGGATTTCTGACCGGCTGGCAATGGAATGGCTCAACAGAGAC 2280
DB 2229 GTGGCTGCTGTGGGGGATTTCTGACCGGCTGGCAATGGAATGGCTCAACAGAGAC 2288
QY 2281 CTCCTCCGATGTCAACAACAGTATCGGGGA CTGTCAATACAGTACAGAGAGGGAGCGCTCC 2340
DB 2289 CTCCTCCGATGTCAACAACAGTATCGGGGA CTGTCAATACAGTACAGAGAGGGAGCGCTCC 2348
QY 2341 CTGGAAGATGTGACCACTTAACTTAATCTTCTGGGAGTTTAAACAGAAACCTCACAGGC 2400
DB 2349 CTGGAAGATGTGACCACTTAACTTAATCTTCTGGGAGTTTAAACAGAAACCTCACAGGC 2408
QY 2401 TAGCGCTGGGCAAGGGTGAATGACAGCAATGCAATTCCTCAATCAAGGCTTCCC 2460

DB 2409 TAGCGCTGGGCAAGGGTGAATGACAGCAATGCAATTCCTCAATCAAGGCTTCCC 2468
QY 2461 GGAGCAAGCTGAGGATCTTATGAGCCAGTCCAACTCTGTATGAGAGAGTGAAGAT 2520
DB 2469 GGAGCAAGCTGAGGATCTTATGAGCCAGTCCAACTCTGTATGAGAGAGTGAAGAT 2528
QY 2521 GATCATAGAGTAAATTTCCATGAGAGACTGAAAGTTCATCAGATGTGATCGAGCAGC 2580
DB 2529 GATCATAGAGTAAATTTCCATGAGAGACTGAAAGTTCATCAGATGTGATCGAGCAGC 2588
QY 2581 GATTTATGACTACTGACAGATTTGAAATCTGTATTTCTGACGCAACTTTGTTGACAT 2640
DB 2589 GATTTATGACTACTGACAGATTTGAAATCTGTATTTCTGACGCAACTTTGTTGACAT 2648
QY 2641 CTCGCAATGCTTGGACGTCCTGTGATAGAGAGTGTCCAGAGTCTGTGTGACCTGTG 2700
DB 2649 CTCGCAATGCTTGGACGTCCTGTGATAGAGAGTGTCCAGAGTCTGTGTGACCTGTG 2708
QY 2701 GACTTCTGAAACCCACTATCATGAGGAGTGTCTGTGGGAAACCCAGAGTGTGACCA 2760
DB 2709 GACTTCTGAAACCCACTATCATGAGGAGTGTCTGTGGGAAACCCAGAGTGTGACCA 2768
QY 2761 GTGCAAGAGGCA----- 2781
DB 2769 GTGCAAGAGGCA----- 2781
QY 2821 CTAGCCAGCTGAGGCTTCAAGCCAGCTACCCAGACAGATGCGGCAAGCTGTGGG 2880
DB 2821 CTAGCCAGCTGAGGCTTCAAGCCAGCTACCCAGACAGATGCGGCAAGCTGTGGG 2888
QY 2881 TCAGGCGGTATGACAGCAGAGAGACTTCTGTGTGTGTGACGCTTCCAGAAC 2940
DB 2881 TCAGGCGGTATGACAGCAGAGAGACTTCTGTGTGTGTGACGCTTCCAGAAC 2948
QY 2941 ATCCAGCTCCTGTCTCGGAGATGAGGCTCCAGATACGTCCTTTTGGCCAGACTGC 3000
DB 2942 ATCCAGCTCCTGTCTCGGAGATGAGGCTCCAGATACGTCCTTTTGGCCAGACTGC 3008
QY 3001 ATCCAGCTCCTGTCTCGGAGATGAGGCTCCAGATACGTCCTTTTGGCCAGACTGC 3060
DB 3001 ATCCAGCTCCTGTCTCGGAGATGAGGCTCCAGATACGTCCTTTTGGCCAGACTGC 3068
QY 3061 GAACCACTTGAAGCAAAACAGAGCCCTGAGACTGAGAGAGAGTGCCTCATGACTGT 3120
DB 3061 GAACCACTTGAAGCAAAACAGAGCCCTGAGACTGAGAGAGAGTGCCTCATGACTGT 3128
QY 3121 CCCACTCAGATGAGGCTTCTGAGAAACCCCTGGAATAGTAACTTACAGCTACCCATC 3180
DB 3121 CCCACTCAGATGAGGCTTCTGAGAAACCCCTGGAATAGTAACTTACAGCTACCCATC 3188
QY 3181 AAGCCAGCAATGAGAACTGGGGCAGTGACTTCTGTATCAGAGTGAAGGCTTCAAT 3240
DB 3181 AAGCCAGCAATGAGAACTGGGGCAGTGACTTCTGTATCAGAGTGAAGGCTTCAAT 3248
QY 3241 AGTGTTCGAACCTCTGTCTCAACAGCTCCGACGAGAGATGCAAAAGTGTGGGCGCTG 3300
DB 3241 AGTGTTCGAACCTCTGTCTCAACAGCTCCGACGAGAGATGCAAAAGTGTGGGCGCTG 3308
QY 3301 GGTGACTCTCTGACTCAGAGAGTGGAGCTGACCAAACTTCCAGTGAAGCTTACCA 3360
DB 3301 GGTGACTCTCTGACTCAGAGAGTGGAGCTGACCAAACTTCCAGTGAAGCTTACCA 3368
QY 3361 TCTTGAAGGGGACTCTTGTGAGCAATGAGGGAATGGAACTTGAAGTCAACACCA 3420
DB 3361 TCTTGAAGGGGACTCTTGTGAGCAATGAGGGAATGGAACTTGAAGTCAACACCA 3428
QY 3421 CTGCGCAATCTGAAAGATTAACCTTACCTCTTGTGCTTCTTCAAGACACTG 3480
DB 3421 CTGCGCAATCTGAAAGATTAACCTTACCTCTTGTGCTTCTTCAAGACACTG 3488
QY 3481 GAGGGGACAGAGGACTTAATGTGAGAGGAGGAGGAGGAGCTGAGGACATGCGCAGCC 3540
DB 3481 GAGGGGACAGAGGACTTAATGTGAGAGGAGGAGGAGGAGCTGAGGACATGCGCAGCC 3548

Dh 3302 GAGGGGACAGCAGAGACTTAAATGTGGCAGCGAAGGGCCAGAGCTAGAGACATGCCAGCC 3361
Qy 3541 CAGGCGCTGGGACCTGTATGAGCGAATGAAAAACAGCCCGACATCAACTGGAGAAAGAC 3600
Dh 3362 CAGGCGCTGGGACCTGTATGAGCGAATGAAAAACAGCCCGACATCAACTGGAGAAAGAC 3421
Qy 3601 TGGAGCTGTGCACTCTTCATTTGGGGTCAACGACTTGTCTATTACTGTAGAAATCCG 3660
Dh 3422 TGGAGCTGTGCACTCTTCATTTGGGGTCAACGACTTGTCTATTACTGTAGAAATCCG 3481
Qy 3661 GAGGCGCTGGGACCTGTATGAGCGAATGAAAAACAGCCCGACATCAACTGGAGAAAGAC 3720
Dh 3482 GAGGCGCTGGGACCTGTATGAGCGAATGAAAAACAGCCCGACATCAACTGGAGAAAGAC 3541
Qy 3721 GAGGAGCTCCCAAGGGCTTCTGTCAAGCTGTGGAGGTCATGAGCTGGAGCTGTAG 3780
Dh 3542 GAGGAGCTCCCAAGGGCTTCTGTCAAGCTGTGGAGGTCATGAGCTGGAGCTGTAG 3601
Qy 3781 CAGGCGCTGGGACCTGTATGAGCGAATGAAAAACAGCCCGACATCAACTGGAGAAAGAC 3840
Dh 3602 CAGGCGCTGGGACCTGTATGAGCGAATGAAAAACAGCCCGACATCAACTGGAGAAAGAC 3661
Qy 3841 CACTCGCAAGCTCTCTGAGAGCAAGAACTGAAGAAAGTGAACCTGCAGCAT 3900
Dh 3662 CACTCGCAAGCTCTCTGAGAGCAAGAACTGAAGAAAGTGAACCTGCAGCAT 3721
Qy 3901 GGCATCTCCAGTTCTCTCTAATGAGCAATGAGCAAGGCTGAGGCTTTGGGGTGTG 3960
Dh 3722 GGCATCTCCAGTTCTCTCTAATGAGCAATGAGCAAGGCTGAGGCTTTGGGGTGTG 3781
Qy 3961 GTGAGCGCTTCTTCCAAAAACACATCACCCTGAGAGAGAGGGGACATGACCTC 4020
Dh 3782 GTGAGCGCTTCTTCCAAAAACACATCACCCTGAGAGAGAGGGGACATGACCTC 3841
Qy 4021 ACCTTCTTCTCCAGAGCTGTTTCACTTCTGAGACCGCGGCTGCGAGATGCCATC 4080
Dh 3842 ACCTTCTTCTCCAGAGCTGTTTCACTTCTGAGACCGCGGCTGCGAGATGCCATC 3901
Qy 4081 GCATCTGGAACAACATGCTGGAACGAGTGGCGGCAAGATCACTCAACAATTGACC 4140
Dh 3902 GCATCTGGAACAACATGCTGGAACGAGTGGCGGCAAGATCACTCAACAATTGACC 3961
Qy 4141 CACAGCGGACCAAACTCAAGTGGCCCTCTCTGAGAGCCCTTCACTCAACCCCTCGG 4200
Dh 3962 CACAGCGGACCAAACTCAAGTGGCCCTCTCTGAGAGCCCTTCACTCAACCCCTCGG 4021
Qy 4201 AACAGCGGATGCTCTCCAGACCAAGCTGGAAGAGCCCCCGAGGTGCTTACTGCGCTGTC 4260
Dh 4022 AACAGCGGATGCTCTCCAGACCAAGCTGGAAGAGCCCCCGAGGTGCTTACTGCGCTGTC 4081
Qy 4261 CCAAGTGGCAGCGGAGTCCGCTTGTGTGGGCAATCTGCGGACAGTGTCTGAGAGTGC 4320
Dh 4082 CCAAGTGGCAGCGGAGTCCGCTTGTGTGGGCAATCTGCGGACAGTGTCTGAGAGTGC 4141
Qy 4321 AGAGAGGTGGCGGAGGAAAGATCTCAATGAGCTGCGACTGCGCTTAG 4377
Dh 4142 AGAGAGGTGGCGGAGGAAAGATCTCTCAATGAGCTGCGACTGCGCTTAG 4198

OS Homo sapiens.
XX
PN W0200272757-A2.
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002W0-US06908.
XX
PR 08-MAR-2001; 2001US-274101P.
PR 08-MAR-2001; 2001US-274194P.
PR 08-MAR-2001; 2001US-274281P.
PR 08-MAR-2001; 2001US-274322P.
PR 09-MAR-2001; 2001US-274849P.
PR 12-MAR-2001; 2001US-275235P.
PR 13-MAR-2001; 2001US-275578P.
PR 13-MAR-2001; 2001US-275579P.
PR 13-MAR-2001; 2001US-275601P.
PR 14-MAR-2001; 2001US-276000P.
PR 16-MAR-2001; 2001US-276776P.
PR 19-MAR-2001; 2001US-276994P.
PR 20-MAR-2001; 2001US-277239P.
PR 20-MAR-2001; 2001US-277321P.
PR 20-MAR-2001; 2001US-277327P.
PR 21-MAR-2001; 2001US-277791P.
PR 22-MAR-2001; 2001US-277833P.
PR 23-MAR-2001; 2001US-278152P.
PR 26-MAR-2001; 2001US-278894P.
PR 27-MAR-2001; 2001US-278999P.
PR 27-MAR-2001; 2001US-279036P.
PR 28-MAR-2001; 2001US-279344P.
PR 30-MAR-2001; 2001US-279388P.
PR 30-MAR-2001; 2001US-279995P.
PR 30-MAR-2001; 2001US-280233P.
PR 02-APR-2001; 2001US-280802P.
PR 02-APR-2001; 2001US-280822P.
PR 02-APR-2001; 2001US-280900P.
PR 04-APR-2001; 2001US-281194P.
PR 13-APR-2001; 2001US-283675P.
PR 30-APR-2001; 2001US-287424P.
PR 02-MAY-2001; 2001US-288066P.
PR 03-MAY-2001; 2001US-288342P.
PR 03-MAY-2001; 2001US-288528P.
PR 15-MAY-2001; 2001US-291190P.
PR 16-MAY-2001; 2001US-291099P.
PR 16-MAY-2001; 2001US-291240P.
PR 30-MAY-2001; 2001US-294485P.
PR 31-MAY-2001; 2001US-294889P.
PR 31-MAY-2001; 2001US-294899P.
PR 18-JUN-2001; 2001US-299027P.
PR 19-JUN-2001; 2001US-299303P.
PR 19-JUN-2001; 2001US-299310P.
PR 10-JUL-2001; 2001US-304354P.
PR 31-JUL-2001; 2001US-309198P.
PR 16-AUG-2001; 2001US-312903P.
PR 10-SEP-2001; 2001US-318462P.
PR 12-SEP-2001; 2001US-318770P.
PR 27-SEP-2001; 2001US-325430P.
PR 27-SEP-2001; 2001US-325681P.
PR 18-OCT-2001; 2001US-330380P.
PR 31-OCT-2001; 2001US-335301P.
PR 14-NOV-2001; 2001US-332172P.
PR 14-NOV-2001; 2001US-332271P.
PR 14-NOV-2001; 2001US-332272P.
PR 14-NOV-2001; 2001US-333184P.
PR 14-NOV-2001; 2001US-333272P.
PR 11-NOV-2001; 2001US-332094P.
PR 03-DEC-2001; 2001US-337426P.
PR 03-DEC-2001; 2001US-338092P.
PR 04-DEC-2001; 2001US-337185P.
PR 03-JAN-2002; 2002US-345705P.
PR 07-MAR-2002; 2002US-0092900.
XX
XX (CURA-) CURAGEN CORP.

Db 1540 CTCACATCTGAGGTTCTCTCGGGCATTTGTGAACTCGGTGACGGTGTGAGATGTCAC 1599
Qy 1681 CTGAGGAGCTGTACAGAGAAAGTCTACTGCCCAAGSAGATCTCAAGTCTCTG 1740
Db 1600 CTGAGGAGCTGTACAGAGAAAGTCTACTGCCCAAGSAGATCTCAAGTCTCTG 1659
Qy 1741 TGTCCCTGTCTCTGAGATTGATGATTAACAAGAACTTGTCTACCTCATGAAATTC 1800
Db 1660 TGTCCCTGTCTCTGAGATTGATGATTAACAAGAACTTGTCTACCTCATGAAATTC 1719
Qy 1801 AACAGAAGTTTACAGAGAAAGCCCACTGATGAGATGGGCAATATGACAAGG 1860
Db 1720 AACAGAAGTTTACAGAGAAAGCCCACTGATGAGATGGGCAATATGACAAGG 1779
Qy 1861 GAAGATTTTACTGTGTGTGAGACGGTCTTGTGAAAACGTGACATGCCAAGAACTCG 1920
Db 1780 GAAGATTTTACTGTGTGTGAGACGGTCTTGTGAAAACGTGACATGCCAAGAACTCG 1839
Qy 1921 GAAGATTTGCTGACACTCTTCTCTGCTCTGACTGTTTCCACTTCAGACAGATCT 1980
Db 1840 GAAGATTTGCTGACACTCTTCTCTGCTCTGACTGTTTCCACTTCAGACAGATCT 1899
Qy 1981 CACTCCGAGCAGCAGCTGCTCTCTGAAACAATATGCTGAGACTGTTGGCCAGAAAGC 2040
Db 1900 CACTCCGAGCAGCAGCTGCTCTCTGAAACAATATGCTGAGACTGTTGGCCAGAAAGC 1959
Qy 2041 ACTGCTATAGTTTGAAGAAACAATATCAATGCTGAGTCCGAAACCGAGTCCA --- GCGC 2097
Db 1960 ACTGCTATAGTTTGAAGAAACAATATCAATGCTGAGTCCGAAACCGAGTCCGAGTCCG 2019
Qy 2098 TTTCTGAGACCTTACAGAAACAGATCAGGGTCTGATGGAGACTGGCTGCCATGAGAGGAC 2157
Db 2020 TTTCTGAGACCTTACAGAAACAGATCAGGGTCTGATGGAGACTGGCTGCCATGAGAGGAC 2079
Qy 2158 AGAGCCCTTCTGCTGCTGACCTTCACTCAGTGCATGCTGAGACCTGCAAGATCTCA 2217
Db 2080 AGAGCCCTTCTGCTGCTGACCTTCACTCAGTGCATGCTGAGACCTGCAAGATCTCA 2139
Qy 2218 GTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2277
Db 2140 GTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2199
Qy 2278 GACTCCCGCATGTACACACAGATACGGGAGCTGTCTATACAGTGCAGAGGGAGCGC 2337
Db 2200 GACTCCCGCATGTACACACAGATACGGGAGCTGTCTATACAGTGCAGAGGGAGCGC 2259
Qy 2338 TCCCTGAGAGATGTGACCACTTAACTTAATATCTTGGGAGTTTAAACAAGAACTGACA 2397
Db 2260 TCCCTGAGAGATGTGACCACTTAACTTAATATCTTGGGAGTTTAAACAAGAACTGACA 2319
Qy 2398 GGGTACGCGTGGGCAAGGGTGTGCAATGACAGCAATGCTTCTCAATCAAGCTGTT 2457
Db 2320 GGGTACGCGTGGGCAAGGGTGTGCAATGACAGCAATGCTTCTCAATCAAGCTGTT 2279
Qy 2458 CCGGAGCAAGGCTGAGATCTTATGAGCAAGTCCAACTCTGATGACAGAAATGAAA 2517
Db 2380 CCGGAGCAAGGCTGAGATCTTATGAGCAAGTCCAACTCTGATGACAGAAATGAAA 2439
Qy 2518 GATGATCATAGATTAATTTTCATGAAAGCTGAAAGTCAATCAAGTGTGATGAGAGC 2577
Db 2440 GATGATCATAGATTAATTTTCATGAAAGCTGAAAGTCAATCAAGTGTGATGAGAGC 2499
Qy 2578 AGGATTTATGATGATCTGACAGATTCGAATGTGATTTCTGAGGCAACTTTGTGAC 2637
Db 2500 AGGATTTATGATGATCTGACAGATTCGAATGTGATTTCTGAGGCAACTTTGTGAC 2559
Qy 2638 CATCTCCGCAATGCTTGAAGCTCTGATAGAGAGGCTGCAAGCTGCTGCTGCAACTC 2697
Db 2560 CATCTCCGCAATGCTTGAAGCTCTGATAGAGAGGCTGCAAGCTGCTGCTGCAACTC 2619
Qy 2698 GTGAGCTTCTGAAACCCCACTATCATGCGAGAGTGTCTCTGAGAAACCCAGCAAGTGC 2757
Db 2620 GTGAGCTTCTGAAACCCCACTATCATGCGAGAGTGTCTCTGAGAAACCCAGCAAGTGC 2679

Qy 2758 CCAATGAGAGGCGCAGCGTTTGTGTAACCTGCTTCTGACCTGCGGAGAACTCCCA 2817
Db 2680 CCAATGAGAGGCGCAGCGTTTGTGTAACCTGCTTCTGACCTGCGGAGAACTCCCA 2739
Qy 2818 GAGCTAGCCAGGCTGAGAGGCTTCAAGCCGAGCTTACCGAGCAGCATGCGAGCTGTG 2877
Db 2740 GAGCTAGCCAGGCTGAGAGGCTTCAAGCCGAGCTTACCGAGCAGCATGCGAGCTGTG 2799
Qy 2878 GGGTCAAGCCGCTATGACAGAGAGAGATTTCTGTGTGTGCTGTGAGCCCTTCTCCAG 2937
Db 2800 GGGTCAAGCCGCTATGACAGAGAGAGATTTCTGTGTGTGCTGTGAGCCCTTCTCCAG 2859
Qy 2938 AACATCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2997
Db 2860 AACATCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2919
Qy 2998 TGCATCCACCCCAATCAGAAATTCACCTCCAGCTGCGCAGAGCCCTTGTGACCAATATG 3057
Db 2920 TGCATCCACCCCAATCAGAAATTCACCTCCAGCTGCGCAGAGCCCTTGTGACCAATATG 2979
Qy 3058 CTGGAACCACTTGTGAAAGCAAAACAGAGACCTGAGACTGAGACAGAGATGCCATCAC 3117
Db 2980 CTGGAACCACTTGTGAAAGCAAAACAGAGACCTGAGACTGAGACAGAGATGCCATCAC 3039
Qy 3118 TGTCCCACTCAGAAATGAGCCCTTCTGAGAAACCCCTCGGAAATGATTAATCAAGTACCC 3177
Db 3040 TGTCCCACTCAGAAATGAGCCCTTCTGAGAAACCCCTCGGAAATGATTAATCAAGTACCC 3099
Qy 3178 ATCAAGCCAGCCATGAGAACTGAGGAGCAGTGAATCTTCTGTATCAGAGTGAAGCTTCC 3237
Db 3100 ATCAAGCCAGCCATGAGAACTGAGGAGCAGTGAATCTTCTGTATCAGAGTGAAGCTTCC 3159
Qy 3238 AATAGTGTTCAACTTGTGTCCACAGCTTCCGACAGCAGACATTCAAAGTGTGTGCGCC 3297
Db 3160 AATAGTGTTCAACTTGTGTCCACAGCTTCCGACAGCAGACATTCAAAGTGTGTGCGCC 3219
Qy 3298 CTGAGTGAATCTTGTGCTGACAGAGTGGGAGCTGACCAAAACAATCCAGTACCTAAC 3357
Db 3220 CTGAGTGAATCTTGTGCTGACAGAGTGGGAGCTGACCAAAACAATCCAGTACCTAAC 3279
Qy 3358 ACATCTTGAAGGAGGACTTCTTGAAGCAATGAGGAGATGAGAACTTGAAGTACACAC 3417
Db 3280 ACATCTTGAAGGAGGACTTCTTGAAGCAATGAGGAGATGAGAACTTGAAGTACACAC 3339
Qy 3418 ACATCTGCCCAACTTGTGAAGAACTTCAACCTTACCTCTTGTGCTTCTTACAGACAC 3477
Db 3340 ACATCTGCCCAACTTGTGAAGAACTTCAACCTTACCTCTTGTGCTTCTTACAGACAC 3399
Qy 3478 TGGAGGAGGAGCAGAGCTTAAATGTGAGCAGGAGGAGGAGCAGAGCTAGAGGACATGCCA 3537
Db 3400 TGGAGGAGGAGCAGAGCTTAAATGTGAGCAGGAGGAGGAGGAGCAGAGCTAGAGGACATGCCA 3459
Qy 3538 GCCAGGCTGAGGAGCTTGTGAGCAGATGAAAGAAACAGGCTCC --- GACATCAACTGAG 3594
Db 3460 GCCAGGCTGAGGAGCTTGTGAGCAGATGAAAGAAAGGCTCCCAAGCACTGAG 3519
Qy 3595 AAAGATGGAAGCTGCTGACACTTCTTATTTGGGGTCAAGACTTGTGTCTTATCTGTGAG 3654
Db 3520 AAAGATGGAAGCTGCTGACACTTCTTATTTGGGGTCAAGACTTGTGTCTTATCTGTGAG 3579
Qy 3655 AATCGGAGGCGCACTTGGCAAGCAATATGTTCAACACATCCAAAGGCGCTGACATC 3714
Db 3580 AATCGGAGGCGCACTTGGCAAGCAATATGTTCAACACATCCAAAGGCGCTGACATC 3627
Qy 3715 CTCTCTGAGAGGCTCCCAAGAGGCTTCTGTCACAGTGTGAGAGTCAATGAGTGTGAGT 3774
Db 3628 CTCTCTGAGAGGCTCCCAAGAGGCTTCTGTCACAGTGTGAGAGTCAATGAGTGTGAGT 3687
Qy 3775 CTGTACAGGCGCAGAGGCGGAAATGTGCATGCTGTGAGCTGACAACTGCACTTGC 3834
Db 3688 CTGTACAGGCGCAGAGGCGGAAATGTGCATGCTGTGAGCTGACAACTGCACTTGC 3747

QY 3835 CTGAGACACTCGAAGAGCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 3894
 DB 3748 CTGAGACACTCGAAGAGCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 3807
 QY 3895 CAGATGAGCACTTCAGATTCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 3954
 DB 3808 CAGATGAGCACTTCAGATTCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 3867
 QY 3955 GTTGAGTGCAGAGCTCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 4014
 DB 3868 GAGACTCGAAGAGCTCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 3924
 QY 4015 GACCTGACCTCTCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 4074
 DB 3925 GACCTGACCTCTCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 3984
 QY 4075 GGCATGACCTCTCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 4134
 DB 3985 GGCATGACCTCTCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 4044
 QY 4135 TTGACCCAGAGAGCAAGAACTGAAAGAACTGAAAGAACTGAACTTC 4194
 DB 4045 TTGACCCAGAGAGCAAGAACTGAAAGAACTGAAAGAACTGAACTTC 4104
 QY 4195 CTGCGGAGAGAGCAAGAACTGAAAGAACTGAAAGAACTGAACTTC 4254
 DB 4105 CTGCGGAGAGAGCAAGAACTGAAAGAACTGAAAGAACTGAACTTC 4164
 QY 4255 GCTGTCCAGAGAGCAAGAACTGAAAGAACTGAAAGAACTGAACTTC 4314
 DB 4165 GCTGTCCAGAGAGCAAGAACTGAAAGAACTGAAAGAACTGAACTTC 4224
 QY 4315 AGGTGAGAGAGAGCAAGAACTGAAAGAACTGAAAGAACTGAACTTC 4374
 DB 4225 AGGTGAGAGAGAGCAAGAACTGAAAGAACTGAAAGAACTGAACTTC 4284
 QY 4375 TAG 4377
 DB 4285 TAG 4287

RESULT 6
 AAD37410 standard; cDNA; 3648 BP.
 AAD37410:
 27-AUG-2002 (first entry)

Human phospholipase-like enzyme encoding cDNA.

Human; phospholipase-like enzyme; cancer; inflammation; Pick's disease; cardiovascular disorder; central nervous system disorder; brain injury; chronic obstructive pulmonary disease; cerebrovascular disease; dementia; Alzheimer's disease; Parkinson's disease; corticobasal degeneration; motor neuron disease; Huntington's disease; Creutzfeldt Jacob dementia; schizophrenia; Korsakoff's psychosis; pain; epilepsy; multiple sclerosis; senile dementia; age associated memory impairment; allergy; asthma; allergic rhinitis; hay fever; atopic dermatitis; cardiovascular disease; anaphylaxis; inflammation; acute respiratory distress syndrome; diabetes; chronic obstructive pulmonary disease; emphysema; obesity; anorexia; overweight; cachexia; bulimia; hypertension; coronary artery disease; type-II diabetes; hyperlipidaemia; gall bladder disease; osteoarthritis; gout; sleep apnoea; respiratory problem; polycystic ovarian syndrome; thrombolytic disease; reduced fertility; pregnancy; stress incontinence; hirsutism; menstrual irregularity; depression; enzyme; gene; ss.

Homo sapiens.

Key Location/Qualifiers
 CDS 1..3648
 /*tag= a
 /product= "Human phospholipase-like enzyme"

PT /note= "CDS does not include start and stop codon"
 FT /partial
 XX
 PN NO200231161-A2.
 PD 18-APR-2002.
 XX
 XX 09-OCT-2001; 2001WO-EP11641.
 PR 10-OCT-2000; 2000US-238445P.
 PR 26-DEC-2000; 2000US-257293P.
 XX
 XX (FARB) BAYER AG.
 PA
 DR WPI; 2002-426287/45.
 DR P-PSDB; AAE22860.
 XX
 PT New human phospholipase-like enzyme polypeptide useful for screening
 PT agents, and in the treatment of cancer, inflammation, diabetes,
 PT obesity, a central nervous system disorder, or a cardiovascular
 PT disorder
 XX
 PS Claim 19; Fig 1; 144p; English.
 XX
 CC The present invention relates to novel human phospholipase-like enzymes
 CC and polynucleotides encoding such proteins. Sequences of the invention
 CC are useful for producing a medicament for modulating the activity of
 CC phospholipase in a disease such as cancer, inflammation, cardiovascular
 CC disorders, chronic obstructive pulmonary diseases, central nervous system
 CC (CNS) disorders such as brain injuries, cerebrovascular disease, dementia
 CC (Alzheimer's disease), Parkinson's disease, corticobasal degeneration,
 CC motor neuron disease, Pick's disease, Huntington's disease, Creutzfeldt
 CC Jacob dementia, schizophrenia with dementia, Korsakoff's psychosis,
 CC pain associated with CNS (e.g. epilepsy, failed back surgery syndrome,
 CC sciatica), multiple sclerosis, stroke, age associated memory impairment,
 CC allergic disease including asthma, allergic rhinitis (hay fever), atopic
 CC dermatitis, anaphylaxis and inflammation, cardiovascular disease, chronic
 CC obstructive pulmonary disease, acute respiratory distress syndrome, gout,
 CC diabetes, emphysema or obesity. They are also used for treating anorexia,
 CC overweight, cachexia, bulimia, hypertension, type-II diabetes, coronary
 CC artery disease, hyperlipidaemia, gall bladder disease, osteoarthritis,
 CC sleep apnoea and respiratory problems, cancer (e.g. breast, prostate,
 CC colon cancer), thrombolytic disease, reduced fertility, polycystic
 CC ovarian syndrome, complications of pregnancy, menstrual irregularity,
 CC hirsutism, stress incontinence and depression. The present sequence is
 CC human phospholipase like enzyme encoding cDNA.
 XX
 SQ Sequence 3648 BP; 899 A; 1032 C; 947 G; 770 T; 0 other;

Query Match 79.1%; Score 3461.2; DB 24; Length 3648;
 Best Local Similarity 97.5%; Pred. No. 0;
 Matches 3558; Conservative 0; Mismatches 33; Indels 57; Gaps 2;

QY 775 GAAGCTGGAACAGCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 834
 DB 1 GAAGCTGGAACAGCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 80
 QY 835 GTTTCAGAGCTCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 894
 DB 61 GTTTCAGAGCTCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 120
 QY 895 CAGATGAGCACTTCAGATTCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 954
 DB 121 CAGATGAGCACTTCAGATTCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 180
 QY 955 AAAGATGAGCACTTCAGATTCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 1014
 DB 181 AAAGATGAGCACTTCAGATTCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTTC 240
 QY 1015 CCTATGTTGAGCAAGAACTGAAAGAACTGAAAGAACTGAAAGAACTGAACTTC 1074

Db 241 CCTATCTGTACGCTACAGAAACGCACTACCTGACAGACTGCAAAACCCCAAGAC 300
Qy 1075 AAGCTT-----AGTTAAGAA 1092
Db 301 AAGCTTGAAGAAATATCATTTTCTTATTTCCCTGTGTGTCTCATTTCAAGTAAAGAA 360
Qy 1093 GAGACGGAATTCAGATGCTCTGACAAAGACCCCTCGATACGGTATCCACTCAGTTTCA 1152
Db 361 GAGACGGAATTCAGATGCTCTGACAAAGACCCCTCGATACGGTATCCACTCAGTTTCA 420
Qy 1153 AGGCTGAAGCCGCTGACATCAATGTAATTTGAGCCCTGGGTACTCTTCAACGGCAGC 1212
Db 421 AGGCTGAAGCCGCTGACATCAATGTAATTTGAGCCCTGGGTACTCTTCAACGGCAGC 480
Qy 1213 AATGAGGCGGGGTCCACACCTGGAGAGTCTTGAAGCTTGAATCTCAAGTACCGAGGCTG 1272
Db 481 AATGAGGCGGGGTCCACACCTGGAGAGTCTTGAAGCTTGAATCTCAAGTACCGAGGCTG 540
Qy 1273 TCCTGAGCGTCCGCGGAGATGAGACATCGGCAACCGTTACACCTGCGGACATCTC 1332
Db 541 TCCTGAGCGTCCGCGGAGATGAGACATCGGCAACCGTTACACCTGCGGACATCTC 600
Qy 1333 CGGGAATTCACCTTCCCTGAGAGGCTTCTCTGTG3CACTGGAAAGAAACCAATCT 1392
Db 601 CGGGAATTCACCTTCCCTGAGAGGCTTCTCTGTG3CACTGGAAAGAAACCAATCT 660
Qy 1393 AATGCTCTTAAACCAAGGCTGTGGCAGAGGCGGAGCTGAGATCACTACCTGTCCAGGC 1452
Db 661 AATGCTCTTAAACCAAGGCTGTGGCAGAGGCGGAGCTGAGATCACTACCTGTCCAGGC 720
Qy 1453 AGGAGCTGTGTGACCTGATGAGAAATGACACGAGATACACTTTCAGAAAGCTGAG 1512
Db 721 AGGAGCTGTGTGACCTGATGAGAAATGACACGAGATACACTTTCAGAAAGCTGAG 780
Qy 1513 AATAAATACCTGTTTAAAGGCGCAATGACTCTGTATTTCTGCAATGATCTGTCCAC 1572
Db 781 AATAAATACCTGTTTAAAGGCGCAATGACTCTGTATTTCTGCAATGATCTGTCCAC 840
Qy 1573 TATTTCTCCAGAACTTCAAGCAATGAGAAAGGCGCTGACATCTTCCATGCTGAG 1632
Db 841 TATTTCTCCAGAACTTCAAGCAATGAGAAAGGCGCTGACATCTTCCATGCTGAG 900
Qy 1633 GTTCTCTCGGCAATTTGTGAACCTGTGACGGTCTTGAATCTTCAACTGAGGAGCTG 1692
Db 901 GTTCTCTCGGCAATTTGTGAACCTGTGACGGTCTTGAATCTTCAACTGAGGAGCTG 960
Qy 1693 TACAGAGAAAGAAAGTCTACCTGCCAAGATGATCTCTAGGTCTGTGTCTCTGTCT 1752
Db 961 TACAGAGAAAGAAAGTCTACCTGCCAAGATGATCTCTAGGTCTGTGTCTCTGTCT 1020
Qy 1753 CTGAAGTTTATGATTAATCTCAAGAACTTGTCTCTCTCATCTGAAATTCACAAAGTTT 1812
Db 1021 CTGAAGTTTATGATTAATCTCAAGAACTTGTCTCTCTCATCTGAAATTCACAAAGTTT 1080
Qy 1813 CAGGAGAGAACCCCACTGATGAGAGTGGGCGATATGACAAAGGAAAGATTTTACT 1872
Db 1081 CAGGAGAGAACCCCACTGATGAGAGTGGGCGATATGACAAAGGAAAGATTTTACT 1140
Qy 1873 GTGCTGTGACGCCGTTCTTTGAAGCGTGAACATGCAAGACCTTGGAAAGATTCCT 1932
Db 1141 GTGCTGTGACGCCGTTCTTTGAAGCGTGAACATGCAAGACCTTGGAAAGATTCCT 1200
Qy 1933 GACAACCTTTCTTCTGCTCTGACTGTTTCCACTTTCAGAGAAAGTCTACCTCCGAGCA 1992
Db 1201 GACAACCTTTCTTCTGCTCTGACTGTTTCCACTTTCAGAGAAAGTCTACCTCCGAGCA 1260
Qy 1993 GCCAGTCTCTCTGAAACAATATGCTGAGGCTGTTGGCGAGAAAGCAACTGCTATAG 2052
Db 1261 GCCAGTCTCTCTGAAACAATATGCTGAGGCTGTTGGCGAGAAAGCAACTGCTATAG 1320
Qy 2053 TTTGAAACAAGATCAATATCAATGTCGAAACCAAGGTCAGCCGTTTCTGAGGACCTAC 2112
Db 1321 TTTGAAACAAGATCAATATCAATGTCGAAACCAAGTTTTCTGTCTCCCTTCAACCTG 1380

Qy 2113 AAGACAG-----CATGACGGGTCAATGAGACCTGGCTGACATGACAGGAC 2157
Db 1381 AGATTCGGGATTTGTTTCTATGCTCTTAGGGGTCAATGAGACCTGGCTGACATGAGGAC 1440
Qy 2158 AGAGCCCTTCTGCTTGAACCTTACCTGACATGACATGCTGAGACCTGACATTCGA 2217
Db 1441 AGAGCCCTTCTGCTTGAACCTTACCTGACATGACATGCTGAGACCTGACATTCGA 1500
Qy 2218 GTTGTGCTGCTTGGGGGATTTCTGACCGCTGGGAAATGGAATTTGGCTTCAACAGAC 2277
Db 1501 GTTGTGCTGCTTGGGGGATTTCTGACCGCTGGGAAATGGAATTTGGCTTCAACAGAC 1560
Qy 2278 GACTCCCCGATTCACACACAGTATCGGGGAGTGTATACAGTACAGGAGGGAGCGC 2337
Db 1561 GACTCCCCGATTCACACACAGTATCGGGGAGTGTATACAGTACAGGAGGGAGCGC 1620
Qy 2338 TCCTGAGAAATGACACACTTACCTAATATCTTTGGGAGTTTAAAGAAACCTTACA 2397
Db 1621 TCCTGAGAAATGACACACTTACCTAATATCTTTGGGAGTTTAAAGAAACCTTACA 1680
Qy 2398 GGTACGCGCTGGGCAAGGCTGATGCCAATGACACGAATGATTTCTCAATCAAGCTGT 2457
Db 1681 GGTACGCGCTGGGCAAGGCTGATGCCAATGACACGAATGATTTCTCAATCAAGCTGT 1740
Qy 2458 CCGGAGCAAAAGGCTGAGATCTTATGAGCAAGTCCAACTCTGATGCAAGATGAAA 2517
Db 1741 CCGGAGCAAAAGGCTGAGATCTTATGAGCAAGTCCAACTCTGATGCAAGATGAAA 1800
Qy 2518 GATGATCATAGATTAATTTTCATGAAAGCTGAAAGCTCATCAAGTCTGATCGAAGC 2577
Db 1801 GATGATCATAGATTAATTTTCATGAAAGCTGAAAGCTCATCAAGTCTGATCGAAGC 1860
Qy 2578 AGGATTTATGATCATCTGACAGATTTGAAATCTGTATTTGACAGCACTTTGTTGAC 2637
Db 1861 AGGATTTATGATCATCTGACAGATTTGAAATCTGTATTTGACAGCACTTTGTTGAC 1920
Qy 2638 CATCTCGCAATGCTTGTGACGCTGTCATGAGAGAGTGGCCAGAGCTGTGCAACTC 2697
Db 1921 CATCTCGCAATGCTTGTGACGCTGTCATGAGAGAGTGGCCAGAGCTGTGCAACTC 1980
Qy 2698 GTGACCTCTGTAACCCCACTATCATGCGGAGAGTTCCTGGGAAACCAAGATGC 2757
Db 1981 GTGACCTCTGTAACCCCACTATCATGCGGAGAGTTCCTGGGAAACCAAGATGC 2040
Qy 2758 CCAATGACAGGCGCAGGCTTTGTGTAACTGTGCTTGAACCTTGGGAGAACTCCCA 2817
Db 2041 CCAATGACAGGCGCAGGCTTTGTGTAACTGTGCTTGAACCTTGGGAGAACTCCCA 2100
Qy 2818 GAGCTAGCCAGGCTGAGGCTTCAAGCCGACCTAACCGAGAGCAATGCGGAGCTGAG 2877
Db 2101 GAGCTAGCCAGGCTGAGGCTTCAAGCCGACCTAACCGAGAGCAATGCGGAGCTGAG 2160
Qy 2878 GGGTCAAGGCGCTATGACACGAGAGAGACTTCTGTGTGTGTGACGCTTCTTCAAG 2937
Db 2161 GGGTCAAGGCGCTATGACACGAGAGAGACTTCTGTGTGTGTGACGCTTCTTCAAG 2220
Qy 2938 AATATCAGTCTCTCTGTCTGTGAGATGGCTTCCAGATGCTGCTTCTTGTGCCCCAGAC 2997
Db 2221 AATATCAGTCTCTCTGTCTGTGAGATGGCTTCCAGATGCTGCTTCTTGTGCCCCAGAC 2280
Qy 2998 TGCATCACCACCAATGAGAAATTCATCTCCAGCTGAGGCGAGGCGCTTGGACCAATAG 3057
Db 2281 TGCATCACCACCAATGAGAAATTCATCTCCAGCTGAGGCGAGGCGCTTGGACCAATAG 2340
Qy 3058 CTGGAACCACTTGGAGCAAAACAGAGACCTTGAACCTTGAACAGAGATGCCATCAC 3117
Db 2341 CTGGAACCACTTGGAGCAAAACAGAGACCTTGAACCTTGAACAGAGATGCCATCAC 2400
Qy 3118 TGTCCACTCAAGATGAGCCTTCTGAGAACCTTGGAAATGTAATCAACGTAACCC 3177
Db 2401 TGTCCACTCAAGATGAGCCTTCTGAGAACCTTGGAAATGTAATCAACGTAACCC 2460

QY 3178 ATCAAGCCAGCATTTGAGAACTGAGGAGTGACTTCTGTGTACAGAGTGAAGGCTTCC 33237
Db 2461 ATCAAGCCAGCATTTGAGAACTGAGGAGTGACTTCTGTGTACAGAGTGAAGGCTTCC 2520
QY 3338 AATTGTGTTCCAACTCTGTCTACACAGTCCGACCAAGACATCAAAAGTGTGGCCG 3297
Db 2521 AATTGTGTTCCAACTCTGTCTACACAGTCCGACCAAGACATCAAAAGTGTGGCCG 2580
QY 3298 CTGGGTAATCTCTGACTGATGAGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGG 3357
Db 2581 CTGGGTAATCTCTGACTGATGAGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGG 2640
QY 3358 ACATCTTGAAGGAGTCTCTTGAAGCATGAGAGGAGTGAAGGAGTGAAGGAGTGAAGG 3417
Db 2641 ACATCTTGAAGGAGTCTCTTGAAGCATGAGAGGAGTGAAGGAGTGAAGGAGTGAAGG 2700
QY 3418 ACATCTGCCCCAATCTTGAAGAGTTCACCTTACTCTTGGCTTCTTACCAAGAC 3477
Db 2701 ACATCTGCCCCAATCTTGAAGAGTTCACCTTACTCTTGGCTTCTTACCAAGAC 2760
QY 3478 TGGAGGGGAGCAGCAGGACTTAATGTGGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 3537
Db 2761 TGGAGGGGAGCAGCAGGACTTAATGTGGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 2820
QY 3538 GCCCAGGCTGGGACTGTGTAGACGATGAAAAAGCCCGACATCAACTGAGAAA 3597
Db 2821 GCCCAGGCTGGGACTGTGTAGACGATGAAAAAGCCCGACATCAACTGAGAAA 2880
QY 3598 GACTGGAAGCTGTGACACTTTCATTGGGGTCAACGACTTGTCTTACTGAGAAAT 3657
Db 2881 GACTGGAAGCTGTGACACTTTCATTGGGGTCAACGACTTGTCTTACTGAGAAAT 2940
QY 3658 CCGGAGGCCCCAATGGGCAAGGATATGTTACAGACATCCAAAGGCCCCGAGACATCTC 3717
Db 2941 CCGGAGGCCCCAATGGGCAAGGATATGTTACAGACATCCAAAGGCCCCGAGACATCTC 3000
QY 3718 TCTGAGAGTCTCCCAAGGAGCTTTCGTCAACGTGTGAGAGTCACTGAGTGAAGCTG 3777
Db 3001 TCTGAGAGTCTCCCAAGGAGCTTTCGTCAACGTGTGAGAGTCACTGAGTGAAGCTG 3060
QY 3778 TACCAGGGCCAAAGGCGGAAAATGTCATGTGGCAGCTCAGAACATCTGACTTGCCTC 3837
Db 3061 TACCAGGGCCAAAGGCGGAAAATGTCATGTGGCAGCTCAGAACATCTGACTTGCCTC 3120
QY 3838 AGACACTCGGAAAGCTCCCTGAGAGACCAAGAACTGAAGAAAGTGAAGTGAAGCTTCCAG 3897
Db 3121 AGACACTCGGAAAGCTCCCTGAGAGACCAAGAACTGAAGAAAGTGAAGTGAAGCTTCCAG 3180
QY 3898 CATGGCATCTCCAGTTCTCTCTACTGGCAACCAATACACAGCGGTGAGGACTTGGGGTT 3957
Db 3181 CATGGCATCTCCAGTTCTCTCTACTGGCAACCAATACACAGCGGTGAGGACTTGGGGTT 3240
QY 3958 GTGGTGAGGCTTTCCTTCAAAAACACTCAACCCACTGAGAGAGAGGAGGAGTGAAG 4017
Db 3241 GTGGTGAGGCTTTCCTTCAAAAACACTCAACCCACTGAGAGAGAGAGGAGGAGTGAAG 3300
QY 4018 CTCACTTCTTCTCCGAGAGTGTTCATCTTCAACCGCGGAGTTCGCCAGATGGCC 4077
Db 3301 CTCACTTCTTCTCCGAGAGTGTTCATCTTCAACCGCGGAGTTCGCCAGATGGCC 3360
QY 4078 ATGCACTCTGGAACAACATGCTGGAACAAGTGGGCGGAAGTACTCTCAACAATTTC 4137
Db 3361 ATGCACTCTGGAACAACATGCTGGAACAAGTGGGCGGAAGTACTCTCAACAATTTC 3420
QY 4138 ACCCAGCGGAGCAAACTCAAGTGCCTCTCTCTGAGAGCCCTTACTCAACCTTC 4197
Db 3421 ACCCAGCGGAGCAAACTCAAGTGCCTCTCTCTGAGAGCCCTTACTCAACCTTC 3480
QY 4198 CGGAAACGCGGATGCTCTCCAGACCAAGCTGGAAGAGCCCCCGAGGTCTCTACTG 4257
Db 3481 CGGAAACGCGGATGCTCTCCAGACCAAGCTGGAAGAGCCCCCGAGGTCTCTACTG 3540
QY 4258 GTCCAGTGGCAGGCGGAGTGGGCTTGTGTGGGAGTCAATCGGAGACAGTGTCTGAGG 4317

Db 3541 GTCCAGTGGCAGGCGGAGTGGGCTTGTGTGGGAGTCAATCGGAGACAGTGTCTGAGG 3600
QY 4318 TGCAGAGAGAGTGGCCCGGAGGAAAGATCTTCCATGAGCTTGGGCACT 4365
Db 3601 TGCAGAGAGAGTGGCCCGGAGGAAAGATCTTCCATGAGCTTGGGCACT 3648
RESULT 7
ABX97048
ID ABX97048 standard; cDNA; 4268 BP.
XX
AC ABX97048;
XX
DT 20-MAY-2003 (first entry)
XX
DE Human NOV24a cDNA.
XX
KM NOVX; cytosolic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200272757-A2.
XX
PD 19-SEP-2002.
XX
PF 08-MAR-2002; 2002MO-US06908.
XX
PR 08-MAR-2001; 2001US-274101P.
PR 08-MAR-2001; 2001US-274194P.
PR 08-MAR-2001; 2001US-274281P.
PR 08-MAR-2001; 2001US-274322P.
PR 09-MAR-2001; 2001US-274849P.
PR 12-MAR-2001; 2001US-275235P.
PR 13-MAR-2001; 2001US-275578P.
PR 13-MAR-2001; 2001US-275579P.
PR 13-MAR-2001; 2001US-275601P.
PR 14-MAR-2001; 2001US-276000P.
PR 16-MAR-2001; 2001US-276776P.
PR 19-MAR-2001; 2001US-276994P.
PR 20-MAR-2001; 2001US-277239P.
PR 20-MAR-2001; 2001US-277321P.
PR 20-MAR-2001; 2001US-277327P.
PR 21-MAR-2001; 2001US-277791P.
PR 22-MAR-2001; 2001US-277833P.
PR 23-MAR-2001; 2001US-278152P.
PR 25-MAR-2001; 2001US-278894P.
PR 27-MAR-2001; 2001US-278999P.
PR 27-MAR-2001; 2001US-279036P.
PR 28-MAR-2001; 2001US-279344P.
PR 30-MAR-2001; 2001US-277338P.
PR 30-MAR-2001; 2001US-279959P.
PR 30-MAR-2001; 2001US-280233P.
PR 02-APR-2001; 2001US-280822P.
PR 02-APR-2001; 2001US-280822P.
PR 02-APR-2001; 2001US-280900P.
PR 04-APR-2001; 2001US-281194P.
PR 13-APR-2001; 2001US-283675P.
PR 30-APR-2001; 2001US-287424P.
PR 02-MAY-2001; 2001US-288066P.
PR 03-MAY-2001; 2001US-288342P.
PR 15-MAY-2001; 2001US-291190P.
PR 15-MAY-2001; 2001US-291099P.
PR 16-MAY-2001; 2001US-291240P.
PR 30-MAY-2001; 2001US-294485P.
PR 31-MAY-2001; 2001US-294889P.
PR 31-MAY-2001; 2001US-294899P.
PR 18-JUN-2001; 2001US-299027P.
PR 19-JUN-2001; 2001US-299303P.
PR 19-JUN-2001; 2001US-299310P.

PR 10-JUL-2001; 2001US-304354P.
PR 31-JUL-2001; 2001US-309198P.
PR 16-AUG-2001; 2001US-312903P.
PR 10-SEP-2001; 2001US-318462P.
PR 12-SEP-2001; 2001US-318770P.
PR 27-SEP-2001; 2001US-325430P.
PR 27-SEP-2001; 2001US-325681P.
PR 18-OCT-2001; 2001US-330380P.
PR 31-OCT-2001; 2001US-335301P.
PR 14-NOV-2001; 2001US-335172P.
PR 14-NOV-2001; 2001US-335271P.
PR 14-NOV-2001; 2001US-335272P.
PR 14-NOV-2001; 2001US-333184P.
PR 14-NOV-2001; 2001US-333272P.
PR 21-NOV-2001; 2001US-333094P.
PR 03-DEC-2001; 2001US-337425P.
PR 03-DEC-2001; 2001US-338092P.
PR 04-DEC-2001; 2001US-337185P.
PR 03-JAN-2002; 2002US-345705P.
PR 07-MAR-2002; 2002US-0092900.
XX
XX (CURA-) CURAGEN CORP.
XX
PI Padigar M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zehushe BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Paturajan M, Gangoli E, Vernet CAM, Guo X, Tchernov V;
PI Fernandez ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;
PI Anderson D, Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H;
PI Alsobrook JP, Lepley DM, Rieger DK;
XX
XX WPI; 2002-723332/78.
DR P-PSDB; ABU65081.
XX
XX NOX polypeptides and polynucleotides, useful for preventing or
PT treating a disorder associated with aberrant NOX expression or
PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
PT bronchial asthma -
XX
XX Claim 13; Page 167-168; 1103pp; English.
XX
XX This invention describes novel human NOX polypeptides which have
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and
CC hypotensive activity. Pharmaceutical compositions comprising the NOX
CC proteins or nucleic acid molecules or NOX antibodies are useful for
CC preventing or treating a disorder associated with aberrant NOX
CC expression or activity e.g. cancer, hypertension, atherosclerosis,
CC cardiomyopathy or bronchial asthma. The products of the invention can
CC be used for gene therapy or in a vaccine. ABX97008-ABX97185 are cDNA
CC fragments amplified and isolated by the PCR primers and probes
CC represented in ABX13460-ABX13462 and ABX97186-ABX97593. ABX97008-ABX97185
CC encode the NOX proteins described in ABU65041-ABU65218.
XX
XX Sequence 4268 BP; 1039 A; 1209 C; 1119 G; 901 T; 0 other;
SQ
Query Match 76.2%; Score 3335; DB 24; Length 4268;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 3512; Conservative 0; Mismatches 50; Indels 57; Gaps 9;
QY 771 TCAGGAAGCTGGAAGCAGCTCTGCTCCAGGAGGTACAGTGAAGAGAGTCTTAC 830
DB 687 TCAGGAAGCTGGAAGCAGCTCTGCTCCAGGAGGTACAGTGAAGAGAGTCTTAC 746
QY 831 CTTGGTTTTCAGAGCTTTCTTATGAGACCACTTCTTACACTGAGAGACCCCG 890
DB 747 CTTGGTTTTCAGAGCTTTCTTATGAGACCACTTCTTACACTGAGAGACCCCG 794
QY 891 ACTCAGAGATTCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 950
DB 795 ACTCAGAGATTCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 854
QY 951 AGAGAAAGATGAGCATTGAGTGTAAACACGAGAGGCAATGAGTGTCTCTTCAGGA 1010
DB 855 AGAGAAAGATGAGCATTGAGTGTAAACACGAGAGGCAATGAGTGTCTCTTCAGGA 914

QY 1011 GAGCCCTATCTGTTGAGTACAGAAACAGCACTACCTGACCTGAGAAACCCCA 1070
DB 915 GAGCCCTATCTGTTGAGTACAGAAACAGCACTACCTGACCTGAGAAACCCCA 974
QY 1071 AGAAGAGTGAAGTAAG 1130
DB 975 AGAAGAGTGAAGTAAG 1091
QY 1131 TACGTTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1190
DB 1032 TACGTTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1091
QY 1191 GGGTGAAGTCTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1250
DB 1092 GGGTGAAGTCTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1151
QY 1251 CTGAGTCACTGAG 1310
DB 1152 CTGAGTCACTGAG 1211
QY 1311 TACCACTGAG 1370
DB 1212 TACCACTGAG 1271
QY 1371 CACTGAG 1430
DB 1272 CACTGAG 1331
QY 1431 TGAGAGTCTACCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1490
DB 1332 TGAG-----CAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1379
QY 1491 ACACCTTCAAG 1550
DB 1380 ACACCTTCAAG 1439
QY 1551 TTTCTGAG 1610
DB 1440 TTTCTGAG 1499
QY 1611 CTTGAG 1664
DB 1500 CTTGAG 1559
QY 1665 GCTTGAAGTGTCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1724
DB 1560 GCTTGAAGTGTCAAGCTGAG 1619
QY 1725 GATCCTCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1784
DB 1620 GATCCTCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1679
QY 1785 TACCTCATGAGATTAACAG 1844
DB 1680 TACCTCATGAGATTAACAG 1739
QY 1845 GCGATATGACACAG 1904
DB 1740 GCGATATGACACAG 1799
QY 1905 CATGCCAAGAACTCGGAG 1964
DB 1800 CATGCCAAGAACTCGGAG 1859
QY 1965 CTTCAAG 2024
DB 1860 CTTCAAG 1919
QY 2025 TGTGGCAG 2084
DB 1920 TGTGGCAG 1979

QY 2085 CCAGGTCACCGCTTTCTGAGGACCTTACAGAACGATGAGGGGTATGAGGACTGGCT 2144
 DB 1980 ACAGGTCACCGCTTTCTGAGGACCTTACAGAACGATGAGGGGTATGAGGACTGGCT 2039
 QY 2145 GCGATGAGGAGACGAGCCCTTTCTGCTTGCACCTTACCTTCAATGATGATGCTGAGACC 2204
 DB 2040 GCGATGAGGAGACGAGCCCTTTCTGCTTGCACCTTACCTTCAATGATGATGCTGAGACC 2099
 QY 2205 TGCAGATCCAACTTGTGCTGCTGAGGGGATTTCTCTGACCGCTGGCAATGGAATTTGG 2264
 DB 2100 TGCAGATCCAACTTGTGCTGCTGAGGGGATTTCTCTGACCGCTGGCAATGGAATTTGG 2159
 QY 2265 CTCGAAACGAGACGACCTCCCGATGTCAACCAACAGATGCGGAGCTGTATACAGTGC 2324
 DB 2160 CTCGAAACGAGACGACCTCCCGATGTCAACCAACAGATGCGGAGCTGTATACAGTGC 2219
 QY 2325 AGGAGGGGAGGCTCCCTGAGGAATGTGACCACTTACCTTAAATCTTCTGGAGTT 2381
 DB 2220 AGGAGGGGAGGCTCCCTGAGGAATGTGACCACTTACCTTAAATCTTCTGGAGTT 2279
 QY 2382 TAAAGAAACCTCAAGGCTTACGCGGTGGGACGCGGTGATGCGCAATGACAGAAATGCAAT 2441
 DB 2280 TAAAGAAACCTCAAGGCTTACGCGGTGGGACGCGGTGATGCGCAATGACAGAAATGCAAT 2339
 QY 2442 CCTCAATCAAGCTGTTCCCGAGGAAAGCTGAGATCTTATGAGCAAGTCCAACTCT 2501
 DB 2340 CCTCAATCAAGCTGTTCCCGAGGAAAGCTGAGATCTTATGAGCAAGTCCAACTCT 2399
 QY 2502 GATGCAAGATGAAAGATGATCATAGATTAATTTCCATGAGAGCTGGAAAGTCAATCAC 2561
 DB 2400 GATGCAAGATGAAAGATGATCATAGATTAATTTCCATGAGAGCTGGAAAGTCAATCAC 2459
 QY 2562 AGTGTATGAGGAGGACGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 2621
 DB 2460 AGTGTATGAGGAGGACGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 2519
 QY 2622 AGCCAACTTTGTTGACCATCTCCGCAATGCTTGGAGAGTCTGATGAGAGTGGCCAG 2681
 DB 2520 AGCCAACTTTGTTGACCATCTCCGCAATGCTTGGAGAGTCTGATGAGAGTGGCCAG 2579
 QY 2682 AGTGTGTGATCCTCGTGTGATCTTCTGAAACCCCACTATCATGCGGAGGTCTGAGG 2741
 DB 2580 AGTGTGTGATCCTCGTGTGATCTTCTGAAACCCCACTATCATGCGGAGGTCTGAGG 2639
 QY 2742 AAACCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2801
 DB 2640 AAACCCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2699
 QY 2802 GCGGAGAACTCCCAAGAGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2861
 DB 2700 GCGGAGAACTCCCAAGAGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2759
 QY 2862 CATGCGGAGCTGGTGGGTGAGGCGCTATGACACGAGAGGAGGAGGAGGAGGAGGAGGAG 2921
 DB 2760 CATGCGGAGCTGGTGGGTGAGGCGCTATGACACGAGAGGAGGAGGAGGAGGAGGAGGAG 2819
 QY 2922 GAGAGCTTTCTTCAAGAACTCCAGCTCTCTGCGGAGTGGGCTCCAGATACGTC 2981
 DB 2820 GAGAGCTTTCTTCAAGAACTCCAGCTCTCTGCGGAGTGGGCTCCAGATACGTC 2879
 QY 2982 CTCTTTGCCCCAGATGATCCACCCCAATTCAGAAATTCACCTCCAGGTGGGAGAGC 3041
 DB 2880 CTCTTTGCCCCAGATGATCCACCCCAATTCAGAAATTCACCTCCAGGTGGGAGAGC 2939
 QY 3042 CCTTTGAGCAATATGCTTGAACCACTTGAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGC 3101
 DB 2940 CCTTTGAGCAATATGCTTGAACCACTTGAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGC 2999
 QY 3102 AGAGATGCCCATACCTGTCCCACTAGATGAGGCTTCTCTGAGAACCTCTGGAGATAG 3161
 DB 3000 AGAGATGCCCATACCTGTCCCACTAGATGAGGCTTCTCTGAGAACCTCTGGAGATAG 3059
 QY 3162 TAACCTACGATCCCATCAAGCAGCATTTGAGAACTGGGGAGTGAATCTTCTGTGTAC 3221

DB 3060 TAACCTACGATCCCATCAAGCAGCATTTGAGAACTGGGGAGTGAATCTTCTGTGTAC 3119
 QY 3222 AGAGTGAAGGCTTCCATATGATGTTCCAACTCTGTCTCCACAGCTCGACACGACAT 3281
 DB 3120 AGAGTGAAGGCTTCCATATGATGTTCCAACTCTGTCTCCACAGCTCGACACGACAT 3179
 QY 3282 CAAAGTGTGGGCGCCCTGGGTACTCTGTGATCAACGAGTGGGAGCTGACCAACAA 3341
 DB 3180 CAAAGTGTGGGCGCCCTGGGTACTCTGTGATCAACGAGTGGGAGCTGACCAACAA 3329
 QY 3342 CTCGAGTACCTTACCCATCTTGAAGGAGCTCTTGGAGATGAGAGGAGATGGGAA 3401
 DB 3240 CTCGAGTACCTTACCCATCTTGAAGGAGCTCTTGGAGATGAGAGGAGATGGGAA 3299
 QY 3402 CTTGAGACTCAACCACTGCGCAACTTCTGAAAGATTTAACCTTACTCTTGG 3461
 DB 3300 CTTGAGACTCAACCACTGCGCAAGTATTTGAAAGATTTCAACCTTACTCTTGG 3359
 QY 3462 CTTCTTACAGGACCTGGGAGGAGGAGCAGCAGGACTAAATGTGGACCGGAGGAGGAGCAG 3521
 DB 3360 CTTCTTACAGGACCTGGGAGGAGGAGCAGCAGGACTAAATGTGGACCGGAGGAGGAGCAG 3419
 QY 3522 AGCT---AGGAGATGCGACGCGGCTGGGACTTGTGAGAGCGAATGAAAAAGAGCCC 3578
 DB 3420 AGCTAGGAGGAGCATGCGACGCGGCTGGGACTTGTGAGAGCGAATGAAAAAGAGCCC 3479
 QY 3579 CGACATCACTGAGAGAAAGCTGAAAGCTGTGACACTTCAATGGGGTCAACGACTT 3638
 DB 3480 ---CATACCTTTCAGAGAAAGCTGAAAGATTAATACCTGTGTAAGGCGGCAATGACT 3536
 QY 3639 GTGTCAATTCGTGAGAAATCCGAGGCGCACTGGGCGACGGAATATGTTCAAGACATCCA 3698
 DB 3537 GTGTCAATTCGTGAGAAATC-----TGTAGGTGAATATGTTCAAGACATCCA 3584
 QY 3699 ACAGGCGCTGAGACATCTCTGTGAGAGCTCCCAAGGCTTTGTGTCAACTGTGAGAGT 3758
 DB 3585 ACAGGCGCTGAGACATCTCTGTGAGAGCTCCCAAGGCTTTGTGTCAACTGTGAGAGT 3644
 QY 3759 CATGAGCTGGCTAGCTGTATCCAGGCGCAAGGCGGAGAAATGTGCCATGTGGACGTCA 3818
 DB 3645 CATGAGCTGGCTAGCTGTATCCAGGCGCAAGGCGGAGAAATGTGCCATGTGGACGTCA 3704
 QY 3819 GAACTGCTGACTTGTGCTGAGACCTCGGAAAGCTCCCTGGAGAAACCAAGAACTGAAAGAA 3878
 DB 3705 GAACTGCTGACTTGTGCTGAGACCTCGGAAAGCTCCCTGGAGAAACCAAGAACTGAAAGAA 3764
 QY 3879 AGTGAACCTGCAAGATGAGATGCAATCTCAAGTTTCTCTTACTGGCAACCAATACACA 3938
 DB 3765 AGTGAACCTGCAAGATGAGATGCAATCTCAAGTTTCTCTTACTGGCAACCAATACACA 3824
 QY 3939 GCGTGAAGACTTGTGCGGTGTGTGTGAGAGCTTTTCTTCAAAACACATCAACCCACTGAA 3998
 DB 3825 GCGTGAAGACTTGTGCGGTGTGTGTGAGAGCTTTTCTTCAAAACACATCAACCCACTGAA 3884
 QY 3999 CGAGAGAGGAGACATGACTCTTCTTCTTCTCGAGAGCTGTTTCACTTCTCAAGACG 4058
 DB 3885 C---AGAGGAGGACATGACTCTTCTTCTTCTCGAGAGCTGTTTCACTTCTCAAGACG 3941
 QY 4059 CGGAGATGCGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4118
 DB 3942 CGGAGATGCGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4001
 QY 4119 GACTTACCTCAACCACTTCAACCAAGCGAGGCAACTCAAGTGGCCCTCTCTGAGAG 4178
 DB 4002 GACTTACCTCAACCACTTCAACCAAGCGAGGCAACTCAAGTGGCCCTCTCTGAGAG 4061
 QY 4179 CCCTTACCTTCAACCTTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4238
 DB 4062 TCCTTACCTTCAACCTTGTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4121
 QY 4239 CGAGGTCTCTACTGGGCTGTCCCAAGTGGAGGAGGAGTGGGCTGTGGTGGGACTCAT 4298

Db 4122 CGAGGCTCTACTGAGGCTGTCCAGTGGCAGCGGAGTGGCCTTGNGTGGGCATCAT 4181
Qy 4299 CGGAGACAGTGTCTGAGGTGCAGAGAGTGGCCGGAGGAGATCCTCCATGAGCCT 4358
Db 4182 CGGAGACAGTGTCTGAGGTGCAGAGAGTGGCCGGAGGAGATCCTCCATGAGCCT 4241
Qy 4359 GCGCACTGTGGCCCTCTAG 4377
Db 4242 GCGCACTGTGGCCCTCTAG 4260

RESULT 8
AAT91874
ID AAT91874 standard; cDNA; 4613 BP.
XX
AC AAT91874;
XX
XX 28-JAN-1998 (first entry)
XX
XX Rat phospholipase-B/lipase cDNA from plasmid pSVL-RiPLB.
XX
XX phospholipase B; lipase; supplement; pancreatic phospholipase; reagent;
XX screening; rat; ds.
XX
XX Rattus rattus.
XX
XX
XX Key Location/Qualifiers
XX CDS 36..4388
XX FT /*tag= a
XX FT sig_peptide 36..125
XX FT /*tag= b
XX FT mat_peptide 126..4385
XX FT /*tag= c
XX FT /note= "putative mature protein"

XX JF09248190-A.
XX
XX
XX PD 22-SEP-1997.
XX
XX PF 15-MAR-1996; 96JP-0086022.
XX
XX PR 15-MAR-1996; 96JP-0086022.
XX
XX PA (TOJO/) TOJO H.
XX PA (TOFU) TONEN CORP.
XX
XX WPI; 1997-520744/48.
XX DR P-PSDB; AAM30751.
XX
XX PT DNA encoding protein with phospholipase B and lipase activity -
XX PT useful as supplement to pancreatic phospholipase
XX
XX PS Claim 4; Pages 7-12; 16pp; Japanese.
XX
XX CC The present cDNA encodes a protein which has both phospholipase B and
XX CC lipase activity. The claimed novel protein comprises at least amino
XX CC acids 367-712 of AAM30751. The phospholipase is useful as a supplement
XX CC to pancreatic phospholipase and as a reagent for the determination of
XX CC phospholipase B/lipase.
XX
XX SQ Sequence 4613 BP; 1218 A; 1243 C; 1142 G; 1010 T; 0 other;

Query Match 61.4%; Score 2687; DB 18; Length 4613;
Best Local Similarity 77.1%; Pred. No. 0;
Matches 3340; Conservative 0; Mismatches 970; Indels 21; Gaps 5;

Qy 28 CTGAGCTGTGCTGCTTCTGGGCAAGGAGCCCTGATGCATACCTCTCTAGAAG 87
Db 69 CTGCTGCTGCTGCTACTGCTGGACAGGAGCCCTCCCAATCCAGGCTCTTCTGAGAA 128
Qy 88 AGTACATTGGAAGGAGCACTATGGCCAGAGACCCTGAAGATTCTCCATTCGATGCAC 147
Db 129 AACCAATCGCAGCCCGACAGCAAGTGTTCGACCTGAAAGATTCTTCAATTCCTTGAAG 188

Qy 148 CCAATTAATTAGAGATGATATGCTTCTTAATTCAGTTCACTCTGAGACCTTGTGAT 207
Db 189 CCAAGAGATTAAATCAGAGTGTCTTCTTAATTCAGTTCACTCTGAGACCTTGAAC 248
Qy 208 ATTAATTTGTGCAAGCATTTGGCAATCTGAAATTCCTTCAGACCCAGGAGCGCG-- 265
Db 249 ATTAATTCGTGGCAGCAGCATCGGCAACTGTAAATCTCTCAGCCCTGCTGAGGGTG 308
Qy 266 ----ATCTGAGAGAGCAAGCTGAGTGAAG--GCCACAGAGAGTGCATGGAGTG 318
Db 309 GTCAACATGAGAGAACTCAAGCCTTCGAGAGCAACTAGAGATGTGCATAGAGATC 368
Qy 319 ATGACAGTCTTTCAGACATCATCAATATTTGATCTCTGTTCCATGCTGTGTGC 378
Db 369 ATGACAGCCTTTCAATATCATCAAGCATTTCAACCTTCTGTTCGAGCCCAAGTGT 428
Qy 379 CACACTGGAAGAGATCATACCCAC--GATGTGTGAGAGACTTGTGATTCAGGCT 435
Db 429 TCTCTGGAGAGGGTACTGAGGCCACACTAATTGCGAGACTTGTGATTCAGGCT 488
Qy 436 CAAGAACTGTGAGAAACATGAAAGAACTGCACTTGAATTGACTGGAAG 495
Db 489 AAAGAGCTGTGAGGATCTGAAAGAACACCCGAACTTGATTTGAGAAAGACTGAAA 548
Qy 496 CTGATCAATGTGTTCTTCAATATGCAAGCCAGTGTTCCTGCTGCTCAACAG 555
Db 549 CTGATCACTGTCTCTTCAATATGCAAGCCAGTGTTCCTGCTGCTCAACAG 608
Qy 556 AATGGCTTGGCGGCGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGAT 615
Db 609 AAAGGCACTGTATGAAGCAGATGAGATGATGATGATGATGATGATGATGATGATGAT 668
Qy 616 GAGTCCCCAGACATTTGTAACCTGTGTGAGACTCTGTGAGTTGACAGAGTCTCTGCT 675
Db 669 GAGTCCCCAGACGTTTGTGAATTTGTGATCTCTGTGAGTTTGAACATGAGCTCAG 728
Qy 676 CAGTATCAGGCACTTGGCTCAGCCCTGCACAGAGCCCTGTAAATGCTCAGAGAGACC 735
Db 729 CAGATCAAGAGACTGTGTTCAAGCCCTGCACAGAGTTTGCAATGCTCAGAGAGATA 788
Qy 736 ACCCGCTGGCCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
Db 789 ACGAAGTTATCCAAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 848
Qy 796 GCTTCCAGAGGATCACTGAGAGAGTCTTCAACGTTGTTTCCAGCTTCTTCTAT 855
Db 849 GCCTCCAGCAAGTTCATTAAGATGAGACCTTGTGTGTTTCCAGTCTTCTTCTCT 908
Qy 856 GAGACCAACCCCATCTCTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 915
Db 909 GAG-----GTAAATCTACCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 962
Qy 916 TGGCATCTCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 975
Db 963 CTGAGATCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1022
Qy 976 AAACACGAGAGCCCAATGAATGTCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1035
Db 1023 GCAAG 1082
Qy 1036 AAGAGCAATCACTGAG 1095
Db 1083 AATAGCAATCACTGAG 1142
Qy 1096 GCGGAATCAGATGTCTGCAAG 1155
Db 1143 ACAAATTTACCTGTCTGCAAG 1202
Qy 1156 CTGAGCGGCTGACATCAAGTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1215
Db 1203 CTGAGCGGCTGACATCAAGTATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1262


```

Db 3423 TCTTGAGCATTTGAGAGGATGGAACGTTGAGAGCCCATACCACTATGCCCAATCTCTG 3482
Qy 3436 AAGAAGTTCAACCCCTTACCTCTTGTGCTTCTTACAGACCTGGAGAGGGAGACAGAGA 3495
Db 3483 AAGAAGTTCAACCCCTTACCTCTTGTGCTTCTTACAGACCCGCTTCCCTGGAGAACAGCAGAGA 3542
Qy 3496 CTAAATGTGACAGGAGAGGGGCGAGAGCTAGGAGACATGCCAGCCAGGCTTGAGACTG 3555
Db 3543 TTAAATGTGACAGAGAGAGGGGCGAGAGCTAGGAGACATGCCAGCCAGGCTTGAGACTG 3602
Qy 3556 GTAGAGGCAATGAAAAACAGCCCGACATCACTTGAGAAAGACTGGAAGCTGTGTACA 3615
Db 3603 GTAGAGGCAATGAAAAACAGCCCTTACCAATCAACATACAGAGAGCTGGAAGCTGTATACA 3662
Qy 3616 CTCTTCATTTGGGGGTCAAGCACTTGTCTTATCTGTAGAGATCCGAGGCCACTTGGCC 3675
Db 3663 CTCTTCATTTGGGGGTCAAGCACTTGTCTTATCTGTAGAGATCCGAGAACCTACTCAACC 3722
Qy 3676 ACGGAATATGTTCAGACATTCACACAGGCGCTGAGACATCTCTTGTAGAGACTCCCAAG 3735
Db 3723 AGGAGATATGTCAAGTATCATTCAGCATGCTTGGACATCTTCTATGAGAGCTTCCAGG 3782
Qy 3736 GCTTTCTCTCAACGCTGTGAGAGCTATGAGAGCTGCTGATCCAGGCGCAAGCGCGG 3795
Db 3783 GTTTTCATCAACGCTGTGAGAGCTATGAGAGCTGCTGATCCAGGCGCGG 3842
Qy 3796 AAATGTGCGCAGAGC---TGAGAGCTCAGAACATGCACTTGCCCTCAGACACTGCGCAAGC 3852
Db 3843 AAATGTGCGCAGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3902
Qy 3853 TCCCTGAGAGAGCAAGAACTGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3912
Db 3903 CTATGAGCAATGAGAGAGCTGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3962
Qy 3913 TTCTCTACTGAGCAACAATACACAGCGTGAGAGACTTTGGCGTTGTGTGACAGCTTTC 3972
Db 3963 CTCTCTACTGAGCAACAATACACAGCGTGAGAGACTTTGGCGTTGTGTGACAGCTTTC 4022
Qy 3973 TTCCAAACACACTCACCCCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4032
Db 4023 TTCCGGAATACCTTGTCCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4082
Qy 4033 GAGAGACTGTTTTCATCTCTGAGACCGGCGGATCCGAGATGAGCATCTGAGACTCTGAGAC 4092
Db 4083 GAGAGACTGTTTTCATCTCTGAGACCGGCGGATCCGAGATGAGCATCTGAGACTCTGAGAC 4142
Qy 4093 AACATGCTGAGAACAGTGCGCGCAAGACTTACCTCCAACTTCCAGCCAGCGGAGCC 4152
Db 4143 AACATGCTGAGAACAGTGCGCGCAAGACTTACCTCCAACTTCCAGCCAGCGGAGCC 4202
Qy 4153 AAATCTAAGTGCGCGCTCTCTGAGAGCCCTTACCTTACACCCCTGCGGAGACAGCCGATG 4212
Db 4203 AAATCTAAGTGCGCGCTCTCTGAGAGCCCTTACCTTACACCTGCTCGGAGACAGCAATG 4262
Qy 4213 CTCCCAAGCAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4272
Db 4263 CTCTTACACAGAGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4322
Qy 4273 GGAAGTCGCGCTTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4332
Db 4323 GTAGAGTGGCTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4382
Qy 4333 CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4343
Db 4383 CAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4393

```

```

DT 16-DEC-2002 (first entry)
XX
DE Human phospholipase protein encoding cDNA SEQ ID NO:1.
XX
KW Human, phospholipase, enzyme, chromosome 2, gene, ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..175
FT CDS 176..1594
FT FT /*tag= a
FT FT /*tag= b
FT FT /product= "phospholipase"
FT FT 1595..1835
FT FT /*tag= c
XX
PN WO200262977-A2.
XX
PD 15-AUG-2002.
XX
PF 28-JUN-2002; 2002MO-US02302.
XX
PR 08-FEB-2001; 2001US-0778961.
XX
PA (PEKE ) PE CORP NY.
XX
PI Van C, Ketchum KA, Di Francesco V, Beasley EM,
XX
DR WPI; 2002-682698/73.
XX
DR P-PSDB; ABP53556.
XX
PT New human phospholipase proteins, useful for the development of human
PT therapeutics and diagnostic compositions, drug screening assays, tissue
PT typing and pharmacogenomic analysis
XX
PS Claim 4; Fig 1A; 95pp; English.
XX
CC The present sequence encodes a human phospholipase protein (1) located on
CC chromosome 2. (1) can be used for identifying agents that modulate its
CC function or activity where the agent is useful for treating a disease
CC or condition mediated by a the human phospholipase protein. (1) peptides
CC can be used in substantial and specific assays related to functional
CC information of the peptide sequences, to raise antibodies or to elicit
CC immune response, as reagents in assays that determine the levels of
CC protein in biological fluids, and as markers for tissues where the
CC corresponding protein is expressed. Nucleotide sequences encoding (1)
CC can be used as probes, primers and chemical intermediates in biological
CC assays, for constructing recombinant vectors, and expressing antigenic
CC portions of the protein. (1) and nucleic acid molecules encoding it can
CC be used in the identification of therapeutic proteins and may serve as
CC models or targets for the development of human therapeutic agents that
CC modulate phospholipase activity in cells and tissues that express the
CC phospholipase, such as in kidney, blood, lung, brain glioblastomas,
CC prostate, colon or leukocytes.
XX
SQ Sequence 1835 BP; 458 A; 549 C; 474 G; 354 T; 0 other;

```

Query Match 31.1%; Score 1362; DB 24; Length 1835;
Best Local Similarity 89.3%; Pred. No. 0;
Matches 1582; Conservative 0; Mismatches 0; Indels 190; Gaps 3;

```

Qy 2618 CTGACGCAACTTTGTGAGACATCTCCGCAATGCTTGAAGCTCCGCAATAGAG----- 2671
Db 1 CTGACGCAACTTTGTGAGACATCTCCGCAATGCTTGAAGCTCCGCAATAGAGACTTT 60
Qy 2672 -----AGTGCCCAAGAGCTGCTGTCACCTCGTGAACCTTCTGAACCCCACTATATGC 2725
Db 61 TCCCTTAGTGAGGCCCAAGTCTCTGTACCTCTGTGAGACTTCTGGAACCCCACTATATGC 120
Qy 2726 GGCAGGTGTTCTTGGAAGAACAGACAGAGTCCCAAGTGACAGAGGCCAGCGTTTGTGTA 2785
Db 121 GGCAGGTGTTCTTGGAAGAACAGACAGAGTCCCAAGTGACAGAGGCCA----- 168

```

QY 2786 ACTGCGTTCTGACCTGCGGGAGAACTCCAGAAGCTAGCCAGGCTGGAGCCTTGAGCC 2845
 Db 169 ----- 168
 QY 2846 GAGCCTAACGGAGACAGATCGCGAGCTGGTGGGGTCAAGCCCTATAGACGACGAGAG 2905
 Db 169 -----GAGCAGCATCGCGAGCTGGTGGGGTCAAGCCCTATAGACGACGAGAG 218
 QY 2906 ACTTCTCTGAGTGTGACAGCCCTTCTTCCAGAACTCCAGCTCCCTGCTGGCGATG 2965
 Db 219 ACTTCTCTGAGTGTGACAGCCCTTCTTCCAGAACTCCAGCTCCCTGCTGGCGATG 273
 QY 2966 GGCTCCCAAGATAGCTCTTCTTTGCCCAAGTGCATCCACCAATCAAGAAATTCACCT 3025
 Db 274 ----- 273
 QY 3026 CCCAGCTGGCAGAGCCCTTTGAGCAATATGCTTGAACCACTTGGAAAGCAAAACAGAGA 3085
 Db 274 -----GCTTGAACCACTTGGAAAGCAAAACAGAGA 302
 QY 3086 CCCTGAGACCTGAGAGAGAGATGCCATCACTGTCTCCACTCAGAAATGAGCCCTTCTGTA 3145
 Db 303 CCCTGAGACCTGAGAGAGAGATGCCATCACTGTCTCCACTCAGAAATGAGCCCTTCTGTA 362
 QY 3146 GAAACCCCTGGAATGATTAATACATGACGTAACCCCAATCAAGCCAGCATTTGAGACTGGGGCA 3205
 Db 363 GAAACCCCTGGAATGATTAATACATGACGTAACCCCAATCAAGCCAGCATTTGAGACTGGGGCA 422
 QY 3206 GTGACTTCTGTGTACAGAGTGAAGGCTTCAATAGTGTTCACCTCTGTCCACGAGC 3265
 Db 423 GTGACTTCTGTGTACAGAGTGAAGGCTTCAATAGTGTTCACCTCTGTCCACGAGC 482
 QY 3266 TCCGACCAAGAGATCAAAAGTGTGCGCCCTGGGTGACTCTTGACTACAGCAGTGG 3325
 Db 483 TCCGACCAAGAGATCAAAAGTGTGCGCCCTGGGTGACTCTTGACTACAGCAGTGG 542
 QY 3326 GAGCTGACCAAACTCCAGTGAACCTCAACCACTTTGGAGGGGACTCTTGGAGCA 3385
 Db 543 GAGCTGACCAAACTCCAGTGAACCTCAACCACTTTGGAGGGGACTCTTGGAGCA 602
 QY 3386 TTGGAGGGGATGGGAACTTGGAGACTCAACCACTGCCCAACTTCTGAAGAGTTCA 3445
 Db 603 TTGGAGGGGATGGGAACTTGGAGACTCAACCACTGCCCAACTTCTGAAGAGTTCA 662
 QY 3446 ACCCTTAACCTCTTGGCTTCTTAACCAAGCACTTGGGAGGGGAGACGACTAATGTGG 3505
 Db 663 ACCCTTAACCTCTTGGCTTCTTAACCAAGCACTTGGGAGGGGAGACGACTAATGTGG 722
 QY 3506 CAGCGGAGGGGAGAGGAGTGAAGCAATGCAAGCCAGAGGCTGGGAGCTGTGAGACGAA 3565
 Db 723 CAGCGGAGGGGAGAGGAGTGAAGCAATGCAAGCCAGAGGCTGGGAGCTGTGAGACGAA 782
 QY 3566 TGAATAAAGAGCCCGACATCAACCTGGAGAAAGACTGGAAGTGTCAACTCTTCAATG 3625
 Db 783 TGAATAAAGAGCCCGACATCAACCTGGAGAAAGACTGGAAGTGTCAACTCTTCAATG 842
 QY 3626 GGGTCAACGACTTGTCTATTACTGTGAGAAATCGAGAGGCCCACTTGGCCAGGAAATATG 3685
 Db 843 GGGTCAACGACTTGTCTATTACTGTGAGAAATCGAGAGGCCCACTTGGCCAGGAAATATG 902
 QY 3686 TTCAGACATTCACAGAGCCCTTGGAGATCTCTCTTAAGAGGCTCCCAAGGGCTTTGTCTA 3745
 Db 903 TTCAGACATTCACAGAGCCCTTGGAGATCTCTCTTAAGAGGCTCCCAAGGGCTTTGTCTA 962
 QY 3746 ACCTGTGAGAGTCAATGAGAGTGGTACCTGTACAGAGGCAAGGGGAAATGTGCCA 3805
 Db 963 ACCTGTGAGAGTCAATGAGAGTGGTACCTGTGTACAGAGGCAAGGGGAAATGTGCCA 1022
 QY 3806 TGTGTGAGAGTCAAGCAACTGCACTTGTCTCAAGCACTTGCAGAAAGCTTCTGGAGAAAGC 3865
 Db 1023 TGTGTGAGAGTCAAGCAACTGCACTTGTCTCAAGCACTTGCAGAAAGCTTCTGGAGAAAGC 1082

QY 3866 AAGAACTGAAGAAAGTGAATGGAACCTTCAGACATGCAATCTCAGTCTTCTTACTGAGC 3925
 Db 1083 AAGAACTGAAGAAAGTGAATGGAACCTTCAGACATGCAATCTCAGTCTTCTTACTGAGC 1142
 QY 3926 ACCAATACACACAGCCTGAGACTTTGGGTTGTGTGACAGCTTTCTTCAAAACACAC 3985
 Db 1143 ACCAATACACACAGCCTGAGACTTTGGGTTGTGTGACAGCTTTCTTCAAAACACAC 1202
 QY 3986 TCACCCCACTGAACGAGAGAGGGGACATGACCTCACTTCTTCTCCGAGGACTGTTTTTC 4045
 Db 1203 TCACCCCACTGAACGAGAGAGGGGACATGACCTCACTTCTTCTCCGAGGACTGTTTTTC 1262
 QY 4046 ACTTCTCAGACCGCGGGAGTGCAGAGATGCCATGCACTCTGGAACCAATGCTGGAAC 4105
 Db 1263 ACTTCTCAGACCGCGGGAGTGCAGAGATGCCATGCACTCTGGAACCAATGCTGGAAC 1322
 QY 4106 CAGTGGCCGCAAGACTACTCTCCACAACTTCAACCCACAGCCGAGCAAACTCAAGTGC 4165
 Db 1323 CAGTGGCCGCAAGACTACTCTCCACAACTTCAACCCACAGCCGAGCAAACTCAAGTGC 1382
 QY 4166 CCTCTCTGAGAGCCCTTACTCTTACACCTGCGGAAAGCCGATGCTCCAGACAGG 4225
 Db 1383 CCTCTCTGAGAGCCCTTACTCTTACACCTGCGGAAAGCCGATGCTCCAGACAGG 1442
 QY 4226 CTGAAGAGCCCCCGAGTGTCTTACTGAGGCTGTCCAGTGGCAGCGGAGTGGCCTTG 4285
 Db 1443 CTGAAGAGCCCCCGAGTGTCTTACTGAGGCTGTCCAGTGGCAGCGGAGTGGCCTTG 1502
 QY 4286 TGTGTGAGTCAATCGGAGAGTGTCTGAGAGTGCAGAGAGGTGGCCGAGGGAAGATC 4345
 Db 1503 TGTGTGAGTCAATCGGAGAGTGTCTGAGAGTGCAGAGAGGTGGCCGAGGGAAGATC 1562
 QY 4346 CTTCATGAGCTGCGGCACTGTGCGCTCTAG 4377
 Db 1563 CTTCATGAGCTGCGGCACTGTGCGCTCTAG 1594

 RESULT 10
 ABX97049
 ID ABX97049 standard; cDNA, 1624 BP.
 XX
 AC ABX97049;
 XX
 DT 20-MAY-2003 (first entry)
 XX
 DE Human NOV24b cDNA.
 XX
 KW NOVX; cytostatic; cardiatic; antiarteriosclerotic; antiasthmatic; cancer;
 KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 KW human; gene; ss.
 OS Homo sapiens.
 XX
 PN MO200272757-A2.
 XX
 PD 19-SEP-2002.
 XX
 XX 08-MAR-2002; 2002MO-US06908.
 PF
 XX 08-MAR-2001; 2001US-274101P.
 PR 08-MAR-2001; 2001US-274194P.
 PR 08-MAR-2001; 2001US-274281P.
 PR 08-MAR-2001; 2001US-274322P.
 PR 09-MAR-2001; 2001US-274849P.
 PR 12-MAR-2001; 2001US-275235P.
 PR 13-MAR-2001; 2001US-275578P.
 PR 13-MAR-2001; 2001US-275579P.
 PR 13-MAR-2001; 2001US-275601P.
 PR 14-MAR-2001; 2001US-276000P.
 PR 16-MAR-2001; 2001US-276776P.
 PR 19-MAR-2001; 2001US-276994P.
 PR 20-MAR-2001; 2001US-277239P.
 PR 20-MAR-2001; 2001US-277321P.

PR 20-MAR-2001; 2001US-277327P.
 PR 21-MAR-2001; 2001US-277791P.
 PR 22-MAR-2001; 2001US-277833P.
 PR 23-MAR-2001; 2001US-278152P.
 PR 25-MAR-2001; 2001US-278899P.
 PR 27-MAR-2001; 2001US-278999P.
 PR 28-MAR-2001; 2001US-279344P.
 PR 30-MAR-2001; 2001US-279388P.
 PR 30-MAR-2001; 2001US-279995P.
 PR 30-MAR-2001; 2001US-280233P.
 PR 02-APR-2001; 2001US-280802P.
 PR 02-APR-2001; 2001US-280822P.
 PR 04-APR-2001; 2001US-280900P.
 PR 13-APR-2001; 2001US-281194P.
 PR 30-APR-2001; 2001US-283675P.
 PR 02-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288342P.
 PR 03-MAY-2001; 2001US-288528P.
 PR 16-MAY-2001; 2001US-291190P.
 PR 16-MAY-2001; 2001US-291099P.
 PR 16-MAY-2001; 2001US-291240P.
 PR 30-MAY-2001; 2001US-294485P.
 PR 31-MAY-2001; 2001US-294889P.
 PR 31-MAY-2001; 2001US-294899P.
 PR 18-JUN-2001; 2001US-299027P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 19-JUN-2001; 2001US-299310P.
 PR 10-JUL-2001; 2001US-304354P.
 PR 31-JUL-2001; 2001US-309198P.
 PR 16-AUG-2001; 2001US-312903P.
 PR 10-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 27-SEP-2001; 2001US-325681P.
 PR 18-OCT-2001; 2001US-330380P.
 PR 31-OCT-2001; 2001US-335301P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-332727P.
 PR 14-NOV-2001; 2001US-333184P.
 PR 21-NOV-2001; 2001US-332029P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 07-MAR-2002; 2002US-0092900.

(CURA-) CURAGEN CORP.

PA Pediaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
 PI Zechusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Palturajan M, Gangoli E, Vernet CAM, Guo X, Tchervnev V;
 PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;
 PI Anderson D, Spederna SK, Catterton E, Burgess C, Lette M, Zhong H;
 PI Alsbrook JP, Lepley DM, Rieger DK;

XX WPI; 2002-723332/78.
 DR P-PSDB; ABU65082.

XX NOVX polypeptides and polynucleotides, useful for preventing or
 PT treating a disorder associated with aberrant NOVX expression or
 PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
 PT bronchial asthma

PS Claim 13; Page 169; 1103pp; English.

CC This invention describes novel human NOVX polypeptides which have
 CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and
 CC hypotensive activity. Pharmaceutical compositions comprising the NOVX
 CC proteins or nucleic acid molecules or NOVX antibodies are useful for

CC preventing or treating a disorder associated with aberrant NOVX
 CC expression or activity e.g. cancer, hypertension, atherosclerosis,
 CC cardiomyopathy or bronchial asthma. The products of the invention can
 CC be used for gene therapy or in a vaccine ABX97008-ABX97185 are cDNA
 CC fragments amplified and isolated by the PCR primers and probes
 CC represented in ABX13460-ABX13462 and ABX97186-ABX97593. ABX97008-ABX97185
 CC encode the NOVX proteins described in ABU65041-ABU65218.

XX Sequence 1624 BP; 422 A; 421 C; 401 G; 380 T; 0 other;

Query Match 26.5%; Score 1160.6; DB 24; Length 1624;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	1161	GGCGGCTGACATAGTAATTGAGCCCTGGTGA	CTCTTCACGGCAGGCAATGGGCG	1220
DB	1	GCCGGCTGACATAGTAATTGAGCCCTGGTGA	CTCTTCACGGCAGGCAATGGGCG	60
QY	1221	CGGGTCCACACCTGGGAGCGTTTGGACGCT	TTGACTGACGAGGCGCTGCTGGAG	1280
DB	61	CGGGTCCACACCTGGGAGCGTTTGGACGCT	TTGACTGACGAGGCGCTGCTGGAG	120
QY	1281	CGTGGCGGAGATGAGAACATCGGCA	CCCTTACCGCCGCAATCTCCGGGAATT	1340
DB	121	CGTGGCGGAGATGAGAACATCGGCA	CCCTTACCGCCGCAATCTCCGGGAATT	180
QY	1341	CAACCCCTCCCTGAAGGGCTTCTCTGTGG	CACTGGGAAAGAACAGCTCTAATGCTT	1400
DB	181	CAACCCCTCCCTGAAGGGCTTCTCTGTGG	CACTGGGAAAGAACAGCTCTAATGCTT	240
QY	1401	CTTAAACAGGCTGTGGCAGAGAGCCGAG	CTGAGATCTA	1460
DB	241	CTTAAACAGGCTGTGGCAGAGAGCCGAG	CTGAGATCTA	300
QY	1461	GGTGAACCTGATGAAGATGACAGAGAT	ACCTTTACAGAAAGACTCGAAGTAATAC	1520
DB	301	GGTGAACCTGATGAAGATGACAGAGAT	ACCTTTACAGAAAGACTCGAAGTAATAC	360
QY	1521	CGGTGTTATAGGGGGAATGACCTCTGTG	ATTTCTGCAATGATCTGATCTCTCC	1580
DB	361	CGGTGTTATAGGGGGAATGACCTCTGTG	ATTTCTGCAATGATCTGATCTCTCC	420
QY	1581	CCAGAACTTACAGACAACTTGAAGAGCC	CTGACATCTCTCATCTGAGGTTCTCG	1640
DB	421	CCAGAACTTACAGACAACTTGAAGAGCC	CTGACATCTCTCATCTGAGGTTCTCG	480
QY	1641	GGCATTGTGAACTGTGTGACGCTGCTG	ATGATGCTCAACCTGAGGAGCTTACACAGA	1700
DB	481	GGCATTGTGAACTGTGTGACGCTGCTG	ATGATGCTCAACCTGAGGAGCTTACACAGA	540
QY	1701	GAAGAAAGTCTATGCGCAAGATGATGAT	CTCAGGTCTGTGCTGCTGCTGCTGCTG	1760
DB	541	GAAGAAAGTCTATGCGCAAGATGATGAT	CTCAGGTCTGTGCTGCTGCTGCTGCTG	600
QY	1761	TGATGATTAATCTCAACAGAACTTGTCT	ACCTCATATGCAATTCACAGAAAGTTTTCAGAGAA	1820
DB	601	TGATGATTAATCTCAACAGAACTTGTCT	ACCTCATATGCAATTCACAGAAAGTTTTCAGAGAA	660
QY	1821	GACCAACCAACTGATTTGAGATGAGTG	CGGATATGACACAGAGGAAGATTTTACTGTGTGT	1880
DB	661	GACCAACCAACTGATTTGAGATGAGTG	CGGATATGACACAGAGGAAGATTTTACTGTGTGT	720
QY	1881	GGAGCGGCTTCTTGAAGAGTGAACATG	CCAAAGACCTCGGAGAGATTCCTGACAACTC	1940
DB	721	GGAGCGGCTTCTTGAAGAGTGAACATG	CCAAAGACCTCGGAGAGATTCCTGACAACTC	780
QY	1941	TTTCTTGCTCTGACTGTTTCCATTTCAG	ACAGCAAGTCTCACTCCGAGACAGCTGTC	2000
DB	781	TTTCTTGCTCTGACTGTTTCCATTTCAG	ACAGCAAGTCTCACTCCGAGACAGCTGTC	840
QY	2001	TCTCTGGAACAATATGCTGGAGCTGTG	GCAGAGACGACTGTCATATAGTTTGAAAA	2060
DB	841	TCTCTGGAACAATATGCTGGAGCTGTG	GCAGAGACGACTGTCATATAGTTTGAAAA	900

QY 2061 CAAGATCAATATCATCATGTCGGAACGAGTCCAGCCGTTTCTGAGAGCTTACAGAAACAG 2120
 Db 901 CAAGATCAATATCATCATGTCGGAACGAGTCCAGCCGTTTCTGAGAGCTTACAGAAACAG 960
 QY 2121 CATGCGAGGTCATGAGGACCTGGGCTCCATGAGGAGAGAGCCCTTCTGCTTGACACC 2180
 Db 961 CATGCGAGGTCATGAGGACCTGGGCTCCATGAGGAGAGAGCCCTTCTGCTTGACACC 1020
 QY 2181 TACCTGAGTGCATGAGCCCTGAGAGCTGAGAGATCAAGTTGTGCTGCTGAGGAGATTC 2240
 Db 1021 TACCTGAGTGCATGAGCCCTGAGAGCTGAGAGATCAAGTTGTGCTGCTGAGGAGATTC 1080
 QY 2241 TCTGACCCGTCGCAATGGAATTGGCTCCAAACAGAGAGACCTCCCGATGTACACACACA 2300
 Db 1081 TCTGACCCGTCGCAATGGAATTGGCTCCAAACAGAGAGACCTCCCGATGTACACACACA 1140
 QY 2301 GTATCGGGAGCTGTCAATACAGTGACAG 2327
 Db 1141 GTATCGGGAGCTGTCAATACAGAGAAAG 1167

RESULT 11

ABA08297 ID ABA08297 standard; cDNA; 802 BP.

XX ABA08297;

XX 11-JAN-2002 (first entry)

XX Human phospholipase B homologue-encoding cDNA, SEQ ID NO:73.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KM haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KM inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KM proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KM myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KM chronic inflammatory condition; proliferative retinopathy;
 KM atherosclerosis; coronary heart disease; arterial ischaemia;
 KM bone disorder; osteoporosis; vascular growth disorder;
 KM tissue regeneration; wound healing; infection; immune disorder;
 KM cell culture; drug screening; gene therapy; antiinflammatory;
 KM antisthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KM cytoskeletal; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KM antifungal; vulnery; antitumor; ss.

XX Homo sapiens.

XX MO200157188-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001MO-US03800.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457740/49.

XX P-PDB; ABB11053.

XX Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -

XX Claim 1; Page 365; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a

CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;

CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation and metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.

XX SQ Sequence 802 BP; 206 A; 236 C; 211 G; 149 T; 0 other;

XX Query Match 15.2%; Score 667.4; DB 22; Length 802;

XX Best Local Similarity 99.9%; Pred. No. 3.5e-177; Indels 0; Gaps 0;

XX Matches 668; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3057 GCTTGAACCACTTGAAGCAAAACAGAGACCTCGACCTGAGAGAGATGCCATCAC 3116
 Db 134 GCTTGAACCACTTGAAGCAAAACAGAGACCTCGACCTGAGAGAGATGCCATCAC 123
 QY 3117 CTGTCCCACTCAGATGAGCCCTTCTGAGAACCCCTCGAATATGTAATCACTGACCC 3176
 Db 194 CTGTCCCACTCAGATGAGCCCTTCTGAGAACCCCTCGAATATGTAATCACTGACCC 253
 QY 3177 CATCAAGCCAGCCATTGAGAACTGGGGCACTGACTTCTGTACAGAGTGAAGCTTC 3236
 Db 254 CATCAAGCCAGCCATTGAGAACTGGGGCACTGACTTCTGTGTACAGAGTGAAGCTTC 313
 QY 3237 CAATAGTGTTCACCACTCTGTCCACAGCAGCTCCGACAGACATCAAGTGTGGCCG 3296
 Db 314 CAATAGTGTTCACCACTCTGTCCACAGCAGCTCCGACAGACATCAAGTGTGGCCG 313
 QY 3297 CTTGGGTGACTCTGTGACTACAGCAGTGGAGCTGACCAAAACAACTCCAGTACCTACC 3356
 Db 374 CTTGGGTGACTCTGTGACTACAGCAGTGGAGCTGACCAAAACAACTCCAGTACCTACC 433
 QY 3357 CACATCTTGAAGAGGAGCTGCTTTGAGAGATTGGAAGGATTTGAGAGCTACAC 3416
 Db 434 CACATCTTGAAGAGGAGCTGCTTTGAGAGATTGGAAGGATTTGAGAGCTACAC 493
 QY 3417 CACACTGCCCAATCTGAAGAGTTCACACCTTACTCTCTGAGCTCTCTACAGAC 3476
 Db 494 CACACTGCCCAATCTGAAGAGTTCACACCTTACTCTCTGAGCTCTCTACAGAC 553
 QY 3477 CTGGAGGGGAGACAGCACTAATGTGCGAGCGAGGGCCAGAGCTTGAAGAGATGCC 3536
 Db 554 CTGGAGGGGAGACAGCACTAATGTGCGAGCGAGGGCCAGAGCTTGAAGAGATGCC 613

QY 3537 AGCCAGGCTGGAGCTGTAGAGCGAATGAAAAAGCCCCGACATCACTGAGAA 3596
 Db 614 AGCCAGGCTGGAGCTGTAGAGCGAATGAAAAAGCCCCGACATCACTGAGAA 673
 QY 3597 AGACTGAAAGCTGTGTCACACTCTTCATTGGGGGTCAAGCACTTGTGTCATTACTGTGAA 3656
 Db 674 AGACTGAAAGCTGTGTCACACTCTTCATTGGGGGTCAAGCACTTGTGTCATTACTGTGAA 733
 QY 3657 TCCGAGAGCCCACTTGGCGACAGAAATATGTTCACACATCCAAAGGCCCTGGACATCCT 3716
 Db 734 TCCGAGAGCCCACTTGGCGACAGAAATATGTTCACACATCCAAAGGCCCTGGACATCCT 793
 QY 3717 CTCTGAGGA 3725
 Db 794 CTCTGAGGA 802
 RESULT 12
 AAH9765 standard; cDNA; 802 BP.
 XX AAH9765;
 AC 16-OCT-2001 (first entry)
 DT 16-OCT-2001 (first entry)
 XX Human protein encoding cDNA sequence SEQ ID NO:600.
 DE Human protein encoding cDNA sequence SEQ ID NO:600.
 XX Human, cancer; ulcer; HIV infection; human immunodeficiency virus;
 KM anti-inflammatory; antirheumatic; antiarthritic; immunosuppressive;
 KM antibacterial; endocrine; cardiac; central nervous system; vitruide;
 KM anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
 KM antiagregant; haemostatic; vulnary; antileuk; osteoporosis; eczema;
 KM dermatological; antiallergic; antiaesthetic; antidiabetic; cystostatic;
 KM neuroprotective; antipresant; nocrotic; antiparkinsonian; infection;
 KM immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KM antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KM cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KM genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KM thrombocytopenia; osteoporosis; severe combined immunodeficiency;
 KM allergic rhinitis; diabetes; multiple sclerosis; depression;
 KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KM neurological disorder; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 XX MO200153455-A2.
 PN 26-JUL-2001.
 PD 22-DEC-2000; 2000MO-US35017.
 PF 23-DEC-1999; 99US-0471275.
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Drmanac RT;
 PI WPI; 2001-457603/49.
 DR P-PSDB; AAM25824.
 XX Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 XX Claim 1; Page 634; 1217p; English.
 XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAM25863. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: anti-inflammatory; antirheumatic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
 CC central nervous system; vitruide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnary;

CC antiulcer; osteopathic; dermatological; antiallergic; antiaesthetic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nocrotic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.
 XX Sequence 802 BP; 206 A; 236 C; 211 G; 149 T; 0 other;
 SQ
 Query Match 15.2%; Score 667.4; DB 22; Length 802;
 Best Local Similarity 99.9%; Pred. No. 3.5e-177;
 Matches 668; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3057 GCTTGAACCACTTGAAGCAAAACAGAGACCTTGAACCTGAGACAGAGATGCCATCAC 3116
 Db 134 GCTTGAACCACTTGAAGCAAAACAGAGACCTTGAACCTGAGACAGAGATGCCATCAC 193
 QY 3117 CTGTCCCACTGAAATGAGAGCCCTTCTGAGAAACCCCTCGGAATATGTAACATCAACCC 3176
 Db 194 CTGTCCCACTGAAATGAGAGCCCTTCTGAGAAACCCCTCGGAATATGTAACATCAACCC 253
 QY 3177 CATCAAGCCAGCACTTGAAGAACTGGGAGAGTACTCTGTGTACAGAGTGAAGGCTTC 3236
 Db 254 CATCAAGCCAGCACTTGAAGAACTGGGAGAGTACTCTGTGTACAGAGTGAAGGCTTC 313
 QY 3237 CAATAGTGTTCACACTCTGTGTCCACAGACTCCGACACAGACATCAAAAGTGTGGCCGC 3296
 Db 314 CAATAGTGTTCACACTCTGTGTCCACAGACTCCGACACAGACATCAAAAGTGTGGCCGC 373
 QY 3297 CTTGGGATCTCTGTACATCAACAGAGTGGAGTTCACCAAAACCTCCAGTACCTAAC 3356
 Db 374 CTTGGGATCTCTGTACATCAACAGAGTGGAGTTCACCAAAACCTCCAGTACCTAAC 433
 QY 3357 CACATCTTGGAGGGGACTCTCTTGGAGCACTTGGAGGGGATGGGAATTGGAGACTCAC 3416
 Db 434 CACATCTTGGAGGGGACTCTCTTGGAGCACTTGGAGGGGATGGGAATTGGAGACTCAC 493
 QY 3417 CACATCTTGGAGGGGACTCTCTTGGAGCACTTGGAGGGGATGGGAATTGGAGACTCAC 3476
 Db 494 CACATCTTGGAGGGGACTCTCTTGGAGCACTTGGAGGGGATGGGAATTGGAGACTCAC 553
 QY 3477 CTGGAGGGGACAGCAGACTAAATGTGGCAGGGGAGGGGCGAGAGCTAGGAGCATGCC 3536
 Db 554 CTGGAGGGGACAGCAGACTAAATGTGGCAGGGGAGGGGCGAGAGCTAGGAGCATGCC 613
 QY 3537 AGCCAGGCTGGAGCTGTGTAAGCAATGAAAAAGCCCCGACATCACTGAGAA 3596
 Db 614 AGCCAGGCTGGAGCTGTGTAAGCAATGAAAAAGCCCCGACATCACTGAGAA 673
 QY 3597 AGACTGAAAGCTGTGTCACACTCTTCATTGGGGGTCAAGCACTTGTGTCATTACTGTGAA 3656
 Db 674 AGACTGAAAGCTGTGTCACACTCTTCATTGGGGGTCAAGCACTTGTGTCATTACTGTGAA 733
 QY 3657 TCCGAGAGCCCACTTGGCGACAGAAATATGTTCACACATCCAAAGGCCCTGGACATCCT 3716
 Db 734 TCCGAGAGCCCACTTGGCGACAGAAATATGTTCACACATCCAAAGGCCCTGGACATCCT 793
 QY 3717 CTCTGAGGA 3725
 Db 794 CTCTGAGGA 802
 RESULT 13
 AAD37412

ID AAD37412 standard; DNA; 608 BP.
 XX AAD37412;
 XX
 DT 27-AUG-2002 (first entry)
 DE Human phospholipase-like enzyme encoding DNA #2.
 XX
 KM Human: phospholipase-like enzyme; cancer; inflammation; Pick's disease;
 KM cardiovascular disorder; central nervous system disorder; brain injury;
 KM chronic obstructive pulmonary disease; cerebrovascular disease; dementia;
 KM Alzheimer's disease; Parkinson's disease; corticobasal degeneration;
 KM motor neuron disease; Huntington's disease; Creutzfeldt Jacob dementia;
 KM schizophrenia; Korsakoff's psychosis; pain; epilepsy; multiple sclerosis;
 KM sciatica; stroke; age associated memory impairment; allergy; asthma;
 KM allergic rhinitis; hay fever; atopic dermatitis; cardiovascular disease;
 KM anaphylaxis; inflammation; acute respiratory distress syndrome; diabetes;
 KM chronic obstructive pulmonary disease; emphysema; obesity; anorexia;
 KM overweight; cachexia; bulimia; hypertension; coronary artery disease;
 KM type-II diabetes; hyperlipidaemia; gall bladder disease; osteoarthritis;
 KM gout; sleep apnoea; respiratory problem; polycystic ovarian syndrome;
 KM thrombolytic disease; reduced fertility; pregnancy; stress incontinence;
 KM hirsutism; menstrual irregularity; depression; enzyme; ds.
 KM
 OS Homo sapiens.
 XX
 PN MO200231161-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 09-OCT-2001; 2001WO-EP11641.
 XX
 PR 10-OCT-2000; 2000US-238445P.
 XX 26-DEC-2000; 2000US-257293P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Zhu Z;
 XX
 DR WP1; 2002-426287/45.
 XX
 PT New human phospholipase-like enzyme polypeptide useful for screening
 PT agents, and in the treatment of cancer, inflammation, diabetes,
 PT obesity, a central nervous system disorder, or a cardiovascular
 PT disorder.
 XX
 PS Disclosure; Fig 5; 144pp; English.
 XX
 CC The present invention relates to novel human phospholipase-like enzymes
 CC and polynucleotides encoding such proteins. Sequences of the invention
 CC are useful for producing a medicament for modulating the activity of
 CC phospholipase in a disease such as cancer, inflammation, cardiovascular
 CC disorders, chronic obstructive pulmonary diseases, central nervous system
 CC (CNS) disorders such as brain injuries, cerebrovascular disease, dementia
 CC (Alzheimer's disease), Parkinson's disease, corticobasal degeneration,
 CC motor neuron disease, Huntington's disease, Creutzfeldt
 CC Jacob dementia, schizophrenia with dementia, Korsakoff's psychosis,
 CC pain associated with CNS (e.g. epilepsy, failed back surgery syndrome,
 CC sciatica), multiple sclerosis, stroke, age associated memory impairment,
 CC allergic disease including asthma, allergic rhinitis (hay fever), atopic
 CC dermatitis, anaphylaxis and inflammation, cardiovascular disease, chronic
 CC obstructive pulmonary disease, acute respiratory distress syndrome, gout,
 CC diabetes, emphysema or obesity. They are also used for treating anorexia,
 CC overweight, cachexia, bulimia, hypertension, type-II diabetes, coronary
 CC artery disease, hyperlipidaemia, gall bladder disease, osteoarthritis,
 CC sleep apnoea and respiratory problems, cancer (e.g. breast, prostate,
 CC colon cancer), thrombolytic disease, reduced fertility, polycystic
 CC ovarian syndrome, complications of pregnancy, menstrual irregularity,
 CC hirsutism, stress incontinence and depression. The present sequence is
 CC human phospholipase like enzyme encoding DNA.
 XX
 SQ Sequence 608 BP; 140 A; 187 C; 158 G; 123 T; 0 other;

Query Match 13.2%; Score 576.2; DB 24; Length 608;
 Best Local Similarity 97.8%; Pred. No. 1,5e-151;
 Matches 584; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 2654 TGGAGCTCTGCTAGAGAGGTGCCAGAGTCTTGCTCAACTGTGACTTCTGAACC 2713
 DB 11 TTGATGGGCTGCATAGAGAGTGTCCAGAGTCTTGCTCAACTGTGACTTCTGAACC 70
 QY 2714 CCAGTATCATGGGGGAGGTGTTCTCTGGGAAACCCAGACAGTCCAGTCCAGAGGCCA 2773
 DB 71 CCACTATCATGGCGGAGGTGTTCTCTGGGAAACCCAGACAGTCCAGTCCAGAGGCCA 130
 QY 2774 GCGTTTGTGTACTGCTGTTCTGACCTCGGAGAACTCCCAAGACTAGCCAGCTGG 2833
 DB 131 GCGTTTGTGTACTGCTGTTCTGACCTCGGAGAACTCCCAAGACTAGCCAGCTGG 190
 QY 2834 AGGCTTTCAGCCGAGCTTACCGGAGAGCATGCGCAGCTGTGGGCTCAGGCGCTATG 2893
 DB 191 AGGCTTTCAGCCGAGCTTACCGGAGAGCATGCGCAGCTGTGGGCTCAGGCGCTATG 250
 QY 2894 ACAAGGAGAGGAGCTTCTGTGTGTGCTGAGCCCTTTCCAGAACATTCAGCTCCCG 2953
 DB 251 ACAAGGAGAGGAGCTTCTGTGTGTGCTGAGCCCTTTCCAGAACATTCAGCTCCCG 310
 QY 2954 TCCTGGCGATGGGCTCCAGATACGTCCTTTTCCCGAGACTGATCCACCAATC 3013
 DB 311 TCCTGGCGATGGGCTCCAGATACGTCCTTTTCCCGAGACTGATCCACCAATC 370
 QY 3014 AGAATTCCACTTCCAGCTGGCCAGAGCCCTTTGACCAATATGCTTGAACCACTTGGAA 3073
 DB 371 AGAATTCCACTTCCAGCTGGCCAGAGCCCTTTGACCAATATGCTTGAACCACTTGGAA 430
 QY 3074 GCAAAACAGAGACCTCTGAGACCTGAGAGCAGAGATGCGCATACCTCTCCACAGAAATG 3133
 DB 431 GCAAAACAGAGACCTCTGAGACCTGAGAGCAGAGATGCGCATACCTCTCCACAGAAATG 490
 QY 3134 AGCCCTTCTGAGAACCCCTCGAATAGTAACTACACGTACCCATCAAGCCAGCATTG 3193
 DB 491 AGCCCTTCTGAGAACCCCTCGAATAGTAACTACACGTACCCATCAAGCCAGCATTG 550
 QY 3194 AGAAGTGGGGCAGTGAATCTTGTGTGACAGAGTGAAGGCTTCCAAATGTTCCAA 3250
 DB 551 AGAAGTGGGGCAGTGAATCTTGTGTGACAGAGTGAAGGCTTCCAAATGTTGTCCAA 607
 RESULT 14
 ID AAL01390
 XX AAL01390 standard; cDNA; 572 BP.
 AC AAL01390;
 DT 21-NOV-2001 (first entry)
 DE Human reproductive system related antigen cDNA SEQ ID NO: 1391.
 XX
 KM Human; reproductive system related antigen; reproductive system disorder;
 KM cancer; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200155320-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01339.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.

PR	19-MAY-2000	2000US-0205515
PR	07-JUN-2000	2000US-0209467
PR	28-JUN-2000	2000US-0214866
PR	30-JUN-2000	2000US-0215135
PR	07-JUL-2000	2000US-0216647
PR	07-JUL-2000	2000US-0216880
PR	11-JUL-2000	2000US-0217487
PR	11-JUL-2000	2000US-0217496
PR	14-JUL-2000	2000US-0218290
PR	26-JUL-2000	2000US-0220963
PR	26-JUL-2000	2000US-0220964
PR	14-AUG-2000	2000US-0222567
PR	14-AUG-2000	2000US-0222567
PR	14-AUG-2000	2000US-0225268
PR	14-AUG-2000	2000US-0225270
PR	14-AUG-2000	2000US-0225447
PR	14-AUG-2000	2000US-0225757
PR	14-AUG-2000	2000US-0225758
PR	14-AUG-2000	2000US-0225759
PR	18-AUG-2000	2000US-0226279
PR	22-AUG-2000	2000US-0226681
PR	22-AUG-2000	2000US-0226882
PR	22-AUG-2000	2000US-0227182
PR	23-AUG-2000	2000US-0227009
PR	30-AUG-2000	2000US-0228924
PR	01-SEP-2000	2000US-0229287
PR	01-SEP-2000	2000US-0229343
PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229509
PR	05-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230438
PR	08-SEP-2000	2000US-0231242
PR	08-SEP-2000	2000US-0231243
PR	08-SEP-2000	2000US-0231244
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0233080
PR	14-SEP-2000	2000US-0233081
PR	12-SEP-2000	2000US-0233196
PR	14-SEP-2000	2000US-0233397
PR	14-SEP-2000	2000US-0233398
PR	14-SEP-2000	2000US-0233399
PR	14-SEP-2000	2000US-0233400
PR	14-SEP-2000	2000US-0233401
PR	14-SEP-2000	2000US-0233403
PR	14-SEP-2000	2000US-0233064
PR	14-SEP-2000	2000US-0233065
PR	21-SEP-2000	2000US-0234223
PR	21-SEP-2000	2000US-0234274
PR	25-SEP-2000	2000US-0234974
PR	25-SEP-2000	2000US-0234988
PR	26-SEP-2000	2000US-0234984
PR	27-SEP-2000	2000US-0235834
PR	27-SEP-2000	2000US-0235837
PR	27-SEP-2000	2000US-0235837
PR	29-SEP-2000	2000US-0236367
PR	29-SEP-2000	2000US-0236368
PR	29-SEP-2000	2000US-0236369
PR	29-SEP-2000	2000US-0236370
PR	02-OCT-2000	2000US-0237037
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237039
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239935
PR	13-OCT-2000	2000US-0239937
PR	20-OCT-2000	2000US-0240960

PR	20-OCT-2000	2000US-0241821
PR	20-OCT-2000	2000US-0241785
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241787
PR	20-OCT-2000	2000US-0241808
PR	20-OCT-2000	2000US-0241809
PR	20-OCT-2000	2000US-0241826
PR	20-OCT-2000	2000US-0246157
PR	08-NOV-2000	2000US-0246417
PR	08-NOV-2000	2000US-0246415
PR	08-NOV-2000	2000US-0246475
PR	08-NOV-2000	2000US-0246476
PR	08-NOV-2000	2000US-0246477
PR	08-NOV-2000	2000US-0246478
PR	08-NOV-2000	2000US-0246523
PR	08-NOV-2000	2000US-0246524
PR	08-NOV-2000	2000US-0246525
PR	08-NOV-2000	2000US-0246526
PR	08-NOV-2000	2000US-0246527
PR	08-NOV-2000	2000US-0246538
PR	08-NOV-2000	2000US-0246532
PR	08-NOV-2000	2000US-0246609
PR	08-NOV-2000	2000US-0246610
PR	08-NOV-2000	2000US-0246611
PR	08-NOV-2000	2000US-0246613
PR	17-NOV-2000	2000US-0249207
PR	17-NOV-2000	2000US-0249208
PR	17-NOV-2000	2000US-0249209
PR	17-NOV-2000	2000US-0249210
PR	17-NOV-2000	2000US-0249211
PR	17-NOV-2000	2000US-0249212
PR	17-NOV-2000	2000US-0249213
PR	17-NOV-2000	2000US-0249214
PR	17-NOV-2000	2000US-0249215
PR	17-NOV-2000	2000US-0249216
PR	17-NOV-2000	2000US-0249217
PR	17-NOV-2000	2000US-0249218
PR	17-NOV-2000	2000US-0249244
PR	17-NOV-2000	2000US-0249245
PR	17-NOV-2000	2000US-0249264
PR	17-NOV-2000	2000US-0249265
PR	17-NOV-2000	2000US-0249297
PR	17-NOV-2000	2000US-0249289
PR	17-NOV-2000	2000US-0249300
PR	01-DEC-2000	2000US-0250160
PR	01-DEC-2000	2000US-0250391
PR	05-DEC-2000	2000US-0251030
PR	05-DEC-2000	2000US-0251988
PR	05-DEC-2000	2000US-0256719
PR	06-DEC-2000	2000US-0251719
PR	08-DEC-2000	2000US-0251856
PR	08-DEC-2000	2000US-0251856
PR	08-DEC-2000	2000US-0251869
PR	08-DEC-2000	2000US-0251989
PR	08-DEC-2000	2000US-0251990
PR	11-DEC-2000	2000US-0254097
PR	05-JAN-2001	2001US-0259678
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX	Rosen CA, Barash SC, Ruben SM	
XX	WPI: 2001-465570/50.	
XX	P-PSDB, AAM95420.	
XX	Claim 1; SEQ ID NO 1391; 1297bp	
XX	Isolated nucleic acid molecule e	
XX	is used in preventing, treating	
XX	The present invention provides t	
XX	number of human reproductive sys	
XX	in the prevention and treatment	
XX	including cancer. The present se	
CC		

Isolated nucleic acid molecule encoding a reproductive system antigen is used in preventing, treating or ameliorating a medical condition -

Claim 1; SEQ ID NO 1391; 1297bp + Sequence Listing; English.

The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a coding sequence of the

CC invention.
XX Sequence 572 BP; 146 A; 159 C; 157 G; 106 T; 4 other;
SQ

Query Match 9.4%; Score 411.2; DB 22; Length 572;
Best Local Similarity 96.7%; Pred. No. 4,6e-105;
Matches 492; Conservative 2; Mismatches 6; Indels 9; Gaps 7;

QY 3042 CCTTTGACCAATATGCTTTGAACCACTTGGAAAGAAAACAGAGACCCCTGGACCTGGAGAGC 3101
DB 15 CCTTTGACCAATATGCTTTGAACCACTTGGAAAGAAAACAGAGACCCCTGGAGAGC 74
QY 3102 AGAGATGCCATCACTGTCTCCACTCAGATGAGCCCTTCTTGAAACCCCTCGAATAG 3161
DB 75 AGAGATGCCATCACTGTCTCCACTCAGATGAGCCCTTCTTGAAACCCCTCGAATAG 134
QY 3162 TAACTACACGTAACCCCATCAAGCCAGCCATTGAGAACTGGGGCAGTGACTTCTGTGTAC 3221
DB 135 TAACTACACGTAACCCCATCAAGCCAGCCATTGAGAACTGGGGCAGTGACTTCTGTGTAC 194
QY 3222 AGAGTGAAGGCTTCAATAGTGTTCACACCTGTGCACAGCTCCGACCGAGAGACAT 3281
DB 195 AGAGTGAAGGCTTCAATAGTGTTCACACCTGTGCACAGCTCCGACCGAGAGACAT 254
QY 3282 CAAAGTGTGCGCCGCTGGGTGACTCTGTGACTACAGAGTGGAGCTCGACCAACAA 3341
DB 255 CAAAGTGTGCGCCGCTGGGTGACTCTGTGACTACAGAGTGGAGCTCGACCAACAA 314
QY 3342 CTCGAGTG-ACTTACCCACATCTTGAAGGGAGCTCTCTT-GGAGCATTGAGGGGATGG 3399
DB 315 CTCGAGTGAACCTTACCCACATCTTGAAGGGAGCTCTTGTGGAGCATTGAGGGGATGG 374
QY 3400 AA--CTTGGAGACTCACACCACTGCCCCAATTTGANAATTCACACCTTACTCC 3457
DB 375 NAACTTGGAGACTCACACCACTGCCCCAATTTGANAATTCACACCTTACTCC 434
QY 3458 TTGGCTTCTCT-ACCAGACCTGGG--AGGGAGACAGAGACT-AAATGTGGACGGAA 3513
DB 435 TTGGCTTCTCTGACACAGACCTGGGGAGGGGAGACAGACGACTGAATGTGGACGGAA 494
QY 3514 GGGGCCAGAGCTA-GGACATGCCAGCCC 3541
DB 495 GGGGCCAGAGTTAGGGGACATGCCAGCCC 523

RESULT 15
ABL96843
ID ABL96843 standard; cDNA; 572 BP.
XX
AC ABL96843;
XX
DT 21-JUN-2002 (first entry)
XX
DE Human testicular antigen encoding cDNA SEQ ID NO: 511.
XX
KW Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cyostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200155317-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01329.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 11-JUL-2000; 2000US-0218230.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226581.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Barash SC, Ruben SM;

XX DR WPI; 2001-483232/52.

XX PT Nucleic acids encoding 973 human testicular antigen polypeptides,
XX useful for preventing, diagnosing and/or treating testicular cancer -

XX PS Claim 1; SEQ ID NO 511, 766bp; English.

XX CC The present invention provides the protein and coding sequences of 973
XX human testicular antigens, and fragments of their genomic sequences. The
XX sequences can be used in the treatment of cardiovascular, urinary system,

CC reproductive system, immune, respiratory, neurological and
CC gastrointestinal disorders, infections, and particularly cancer,
CC especially testicular cancers. The present sequence is a cDNA of the
CC invention.

XX Sequence 572 BP; 146 A; 159 C; 157 G; 106 T; 4 other;

Query Match 9.4%; Score 411.2; DB 23; Length 572;
Best Local Similarity 96.7%; Pred. No. 4,6e-105;
Matches 492; Conservative 2; Mismatches 6; Indels 9; Gaps 7;

QY 3042 CCTTGGACCAATATGCTTGAACCACTTGGAAACAAAGAGACCTTGGACCTGAGAGC 3101
DB 15 CTTTGGACCAATATGCTTGAACCACTTGGAAACAAAGAGACCTTGGACCTGAGAGC 74

QY 3102 AGAGATGCCCATACCTGCTCCCACTAGATGAGCCCTTCTTGAAGAACCTTGGAAATAG 3161
DB 75 AGAGATGCCCATACCTGCTCCCACTAGATGAGCCCTTCTTGAAGAACCTTGGAAATAG 134

QY 3162 TAACTACAGTACCCCATCAAGCCAGCCAGCATTTGAACTGGGGCAGTACTTCTGTGTAC 3221
DB 135 TAACTACAGTACCCCATCAAGCCAGCCAGCATTTGAACTGGGGCAGTACTTCTGTGTAC 194

QY 3222 AGAGTGAAGGCTTCCATATGCTTCCAACTCTGTCCACAGCTCCGACAGACAT 3281
DB 195 AGAGTGAAGGCTTCCATATGCTTCCAACTCTGTCCACAGCTCCGACAGACAT 254

QY 3282 CAAAGTGTGCGCCCTCGGTGATCTCTGATACAGCGTGGAGCTTGACCAACAA 3341
DB 255 CAAAGTGTGCGCCCTCGGTGATCTCTGATACAGCGTGGAGCTTGACCAACAA 314

QY 3342 CTCGAGTG-ACCTACCAATCTTGGAGGGGACTCTCTT-GGAGCATTTGGAGGGATGGG 3399
DB 315 CTCGAGTGAACTTACCAATCTTGGAGGGGACTCTTGGAGCATTTGGAGGGATGGG 374

QY 3400 AA--CTTGAAGACTACACCACTGCCCCCAATTTGAAAGATTCAACCTTACCTCC 3457
DB 375 NAACTTGGAGACTCAACACACTGCCCAATTTGAAAGATTCAACCTTACCTCC 434

QY 3458 TTGGCTTCTCT-ACGACGACTTGGG--AGGGAGCAGCAGACT-AAATGTGGAGCGGAA 3513
DB 435 TTGGCTTCTCTGACGACGACTTGGGAGGGGAGCAGCAGACTGAAATGTGGAGCGGAA 494

QY 3514 GGGGCCAGAGCTA-GGAGCATGCGAGCCC 3541
DB 495 GGGGCCAGAGTTAGGGGACATGCCAGCCC 523

Search completed: January 6, 2004, 08:00:25
Job time : 1748 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 04:33:27 : Search time 15537 Seconds

(without alignments)
11524.840 Million cell updates/sec

Title: US-10-054-691-1

Perfect score: 4377
Sequence: 1 atcgagcctgcgcagcagcat.....tgcgacactgctgcctctcag 4377

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

GenBank1:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_ph:*
23: em_pat:*
24: em_pl:*
25: em_ro:*
26: em_sbs:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_in:*
32: em_hcg_om:*
33: em_hcg_ov:*
34: em_hcg_pat:*
35: em_hcg_ph:*
36: em_hcg_pl:*
37: em_hcg_ro:*
38: em_hcg_sy:*
39: em_hcg_un:*
40: em_hcg_vl:*
41: em_hcg_hum:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4377	100.0	4377	6	AX492941
2	3461.2	79.1	3648	6	AX417817
3	3124	71.4	4612	4	OCPRHLP
4	2962	67.7	4508	10	AF045454
5	2687	61.4	4590	10	D63648
6	2687	61.4	4513	6	E13935
7	1362	31.1	1835	6	AX711962
8	1319.4	30.1	1790	6	BC042674
9	970	22.2	2477	6	AX713706
10	970	22.2	2477	6	AK055428
11	576.2	13.2	608	6	AX417821
12	372.2	8.5	583	6	AX417820
13	322.2	7.4	820	10	BC033606
14	251	5.7	265	6	AX417822
15	237.4	5.4	446	6	AX417823
16	205	4.7	43543	6	AX711964
17	205	4.7	180465	9	AC074011
18	177.2	4.0	163595	2	AC022076
19	120.6	2.8	1672	3	AY069174
20	109	2.5	103926	9	AC093164
21	107.8	2.5	357	6	AX314612
22	107.2	2.4	50250	2	AC017341
23	107.2	2.4	179283	3	AC009383
24	107.2	2.4	299903	3	AB003514
25	94.6	2.2	101340	2	AC018104
26	94.6	2.2	190668	3	AC007419
27	94.6	2.2	260731	3	AB003611
28	85.6	2.0	217253	2	AC102372
29	82.8	1.9	214038	2	AC141490
30	81.6	1.9	214038	2	AC141490
31	75.6	1.7	217253	2	AC102372
32	70.4	1.6	134246	9	AC125617
33	68.6	1.6	684	4	CFA243475
34	68.2	1.6	125020	9	AF429315
35	63.2	1.4	125020	9	AF429315
36	61.6	1.4	631	11	BV055548
37	59.8	1.4	44237	3	AC024847
38	59.8	1.4	267118	2	AC006889
39	58	1.3	198736	5	AL929023
40	52.6	1.2	2000	6	AX655393
41	52.6	1.2	123789	2	AC110674
42	52.2	1.2	2000	6	AX655393
43	49.2	1.1	232304	2	AC111030
44	47.8	1.1	34305	3	CEW02812
45	47.2	1.1	212127	5	BX248515

ALIGNMENTS

RESULT 1	AX492941	4377 bp	DNA	linear	PAT 26-SEP-2002
LOCUS	AX492941				
DEFINITION	Sequence 1 from Patent WO02059328.				
ACCESSION	AX492941				
VERSION	AX492941.1				
KEYWORDS	GI:23338611				
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE					
AUTHORS	Turner Jr,C.A., Miranda,M. and Yu,X.				
TITLE	Human lipase and polynucleotides encoding the same				
JOURNAL	Patent: WO 02059328-A 1 01-AUG-2002;				

LEXICON GENETICS INC (US)
 Location/Qualifiers
 source 1..4377
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 1085 a 1217 c 1155 g 920 t
 ORIGIN

Query Match 100.0%; Score 4377; DB 6; Length 4377;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGGGGGTGGGCGAGGAGGATTTCTCTGAGAGTGGTGGTCTTCTGGGCAAGGAGC 60
DB 1 ATGGGGGTGGGCGAGGAGGATTTCTCTGAGAGTGGTGGTCTTCTGGGCAAGGAGC 60
QY 61 CCTCAGATCCATCTCTCTAGAAAGATACATTTGAGAGGCGAGCTATGCGCAGAGACC 120
DB 61 CCTCAGATCCATCTCTCTAGAAAGATACATTTGAGAGGCGAGCTATGCGCAGAGACC 120
QY 121 CTGAAGAAATTCCTCATTCCTCATGCAACCAATTAATAGAGTGAATATGCTTTCTAAA 180
DB 121 CTGAAGAAATTCCTCATTCCTCATGCAACCAATTAATAGAGTGAATATGCTTTCTAAA 180
QY 181 TCAGTTCATCTCTGAGAGCTTCTGATTAATTTGTCAGAGCCATTTGGCAATCTGAA 240
DB 181 TCAGTTCATCTCTGAGAGCTTCTGATTAATTTGTCAGAGCCATTTGGCAATCTGAA 240
QY 241 ATTCTCTCAGACCCAGGAGCGGGGATCTGAGAGAGCAAGCTGAGCTGAAGGCCAG 300
DB 241 ATTCTCTCAGACCCAGGAGCGGGGATCTGAGAGAGCAAGCTGAGCTGAAGGCCAG 300
QY 301 CAGGTGCGATGGAGAGTGAAGAGTCTTTCAAGATCATCAATATTTCAAGTCTTCT 360
DB 301 CAGGTGCGATGGAGAGTGAAGAGTCTTTCAAGATCATCAATATTTCAAGTCTTCT 360
QY 361 GTTCCAAATGCTGTGTGTCACACTGAAAGAGAGTCAATCCCAAGATGGTCTGAAGAC 420
DB 361 GTTCCAAATGCTGTGTGTCACACTGAAAGAGAGTCAATCCCAAGATGGTCTGAAGAC 420
QY 421 TTGTGATTCAGGCTCAAGACTGTGTGAGAAACATGAAGAGAACCTGCAACTTGA 480
DB 421 TTGTGATTCAGGCTCAAGACTGTGTGAGAAACATGAAGAGAACCTGCAACTTGA 480
QY 481 CAATTTGACCTGGAAGCTCATCATGTGTCTTCAATGATGCAAGCCAGTGTACCTGTGC 540
DB 481 CAATTTGACCTGGAAGCTCATCATGTGTCTTCAATGATGCAAGCCAGTGTGTACCTGTGC 540
QY 541 CCTCTGCTCAACAGAAATGGGCTTGGCGGGCGGCGGTGATGAGCTGATGGGGGTCTG 600
DB 541 CCTCTGCTCAACAGAAATGGGCTTGGCGGGCGGCGGTGATGAGCTGATGGGGGTCTG 600
QY 601 GACTACCTGAGAGAGAGGTCCCAAGACATTTGTAAACCTGTGTGAAGCTCTCTGAGGT 660
DB 601 GACTACCTGAGAGAGAGGTCCCAAGACATTTGTAAACCTGTGTGAAGCTCTCTGAGGT 660
QY 661 GCGAGAGTCTCTGTCTGATACAGGCACTTGGCTCAAGCCCTGCAAGAGCCCTGTAAAT 720
DB 661 GCGAGAGTCTCTGTCTGATACAGGCACTTGGCTCAAGCCCTGCAAGAGCCCTGTAAAT 720
QY 721 TGGTCAGAGAGAGACACCGGCTGGGCAAGTGTGTATGACAGTGTCTTATCAGAGAGCC 780
DB 721 TGGTCAGAGAGAGACACCGGCTGGGCAAGTGTGTATGACAGTGTCTTATCAGAGAGCC 780
QY 781 TGGAAACAGCTCTGCTGCTCAGAGAGTACAGTGAAGAGAGTCTTCAACGTTGTTTC 840
DB 781 TGGAAACAGCTCTGCTGCTCAGAGAGTACAGTGAAGAGAGTCTTCAACGTTGTTTC 840
QY 841 CAGGCTTTCTTCTATGAGACCAACCTCTTACACTGAGAGAGCCCGCACTTCAGAGAT 900
DB 841 CAGGCTTTCTTCTATGAGACCAACCTCTTACACTGAGAGAGCCCGCACTTCAGAGAT 900

```

```

QY 901 TTACACAGCTGGCTGGCATCTCTGGAATAGATGATGAGAGCCAGAGAGAGAAAGAT 960
DB 901 TTACACAGCTGGCTGGCATCTCTGGAATAGATGATGAGAGCCAGAGAGAGAAAGAT 960
QY 961 GAGCCATTGAGTGTAAACACGGGAGGCCATGAGTGTCTCTCAGAGAGAGCCCTTAT 1020
DB 961 GAGCCATTGAGTGTAAACACGGGAGGCCATGAGTGTCTCTCAGAGAGAGCCCTTAT 1020
QY 1021 CTGTGAGCTAGAGAAACAGCACTACCTGACAGCTGACAGAAACCCCAAGCAAGCTT 1080
DB 1021 CTGTGAGCTAGAGAAACAGCACTACCTGACAGCTGACAGAAACCCCAAGCAAGCTT 1080
QY 1081 GAGTAAAGAGAGAGACCGGAATCAGATGTCTGACAAAGACCCCTCCGATACGTTCCC 1140
DB 1081 GAGTAAAGAGAGAGACCGGAATCAGATGTCTGACAAAGACCCCTCCGATACGTTCCC 1140
QY 1141 ACCTCAGTTCAATGCTGAAAGCCGCTGACATCAAGTAAATGAGAGCCCTGGTACTCT 1200
DB 1141 ACCTCAGTTCAATGCTGAAAGCCGCTGACATCAAGTAAATGAGAGCCCTGGTACTCT 1200
QY 1201 CTGACGGCAGGCAATGGGGCCGGGTCACACCTGGGAAAGCTTTGACGCTTGAACCTCAG 1260
DB 1201 CTGACGGCAGGCAATGGGGCCGGGTCACACCTGGGAAAGCTTTGACGCTTGAACCTCAG 1260
QY 1261 TACCGAGGCTGTCTCTGAGAGCTGTGGCGAGATGAGAACATCGGCAACCGTTACACCTTG 1320
DB 1261 TACCGAGGCTGTCTCTGAGAGCTGTGGCGAGATGAGAACATCGGCAACCGTTACACCTTG 1320
QY 1321 GCGAATCTCTCGGGAATTCACCTTCTCTGAAAGGCTTCTCTGTGGCACTGGGAAA 1380
DB 1321 GCGAATCTCTCGGGAATTCACCTTCTCTGAAAGGCTTCTCTGTGGCACTGGGAAA 1380
QY 1381 GAAACAGTCTTAATGCTTTTAAACAGGCTGTGGCAGAGGCGGAGCTGAGATCTTA 1440
DB 1381 GAAACAGTCTTAATGCTTTTAAACAGGCTGTGGCAGAGGCGGAGCTGAGATCTTA 1440
QY 1441 CCTGTCAGGCGCAGAGGCTGTGTGAAGCTGATGAAGATGACACGAGATACACTTTCAG 1500
DB 1441 CCTGTCAGGCGCAGAGGCTGTGTGAAGCTGATGAAGATGACACGAGATACACTTTCAG 1500
QY 1501 GAAGACTGGAAGATTAACCTGTTTATAGCGGCAATGACCTGTGTGATTTCTGCAAT 1560
DB 1501 GAAGACTGGAAGATTAACCTGTTTATAGCGGCAATGACCTGTGTGATTTCTGCAAT 1560
QY 1561 GATCTGTCCACTATCTCCCAAGAACTTCAAGAGAAACATTTGAGAAAGCCCTGAGCATC 1620
DB 1561 GATCTGTCCACTATCTCCCAAGAACTTCAAGAGAAACATTTGAGAAAGCCCTGAGCATC 1620
QY 1621 CTCATGCTGAGGTTCTCTCGGCAATTTGTGAACCTGTGTGACGCTGTGAGATCGTCAAC 1680
DB 1621 CTCATGCTGAGGTTCTCTCGGCAATTTGTGAACCTGTGTGACGCTGTGAGATCGTCAAC 1680
QY 1681 CTGAGGAGCTGTACAGAGAGAAAGATCTACTGCCAAGATGATCTCAGGTCTCTG 1740
DB 1681 CTGAGGAGCTGTACAGAGAGAAAGATCTACTGCCAAGATGATCTCAGGTCTCTG 1740
QY 1741 TGTCCCTGTGTCTGGAAGTTTGTATGATTAATCAACGAAGTCTGACCTCATGCAATTC 1800
DB 1741 TGTCCCTGTGTCTGGAAGTTTGTATGATTAATCAACGAAGTCTGACCTCATGCAATTC 1800
QY 1801 AACAGAAGTTTCAAGAGAAAGACCAACCACTGATTTGAGAGTGGGAGATGACAAAG 1860
DB 1801 AACAGAAGTTTCAAGAGAAAGACCAACCACTGATTTGAGAGTGGGAGATGACAAAG 1860
QY 1861 GAAGATTTTACTGTGTGTGACAGCCGTTCTTTGAAAACGTGACATGCCAAAGACCTCG 1920
DB 1861 GAAGATTTTACTGTGTGTGACAGCCGTTCTTTGAAAACGTGACATGCCAAAGACCTCG 1920
QY 1921 GAAGATTTGCTGACAACTTTCTTCTGCTCTGACATGTTTCCACTTCAGACCAAGTCT 1980
DB 1921 GAAGATTTGCTGACAACTTTCTTCTGCTCTGACATGTTTCCACTTCAGACCAAGTCT 1980
QY 1981 CACTCCGAGACAGAGTGTCTCTGGAACAAATATGCTGAGAGCTGTGGCCAGAGAGC 2040

```

```
Db 1981 GACTCCGAGCAGCCAGTGTCTCTGGAAACAATATGTGGAGCTGTGGCCAGAAAGACG 2040
Qy 2041 ACTCGTCATAGTTTGAAGAAACAATCAATATCAATATGTCGGAACAAGGTCCAGCCGTTT 2100
Db 2041 ACTCGTCATAGTTTGAAGAAACAATCAATATCAATATGTCGGAACAAGGTCCAGCCGTTT 2100
Qy 2101 CTGAGGACCTTACAGAAACAGATGACAGGGTCAATGGGACCTGGTGCATGACAGGACAGA 2160
Db 2101 CTGAGGACCTTACAGAAACAGATGACAGGGTCAATGGGACCTGGTGCATGACAGGACAGA 2160
Qy 2161 GCGCCCTTCTGCTTGGCACTCTTACCTGATGTCAGTCCCTGAGACCTTGACATATCAAGTT 2220
Db 2161 GCGCCCTTCTGCTTGGCACTCTTACCTGATGTCAGTCCCTGAGACCTTGACATATCAAGTT 2220
Qy 2221 GTGGCTGCTTGGGGGATTTCTGACCGCTGGGCAATGGAAATTTGGCTCCAAACAGACGAC 2280
Db 2221 GTGGCTGCTTGGGGGATTTCTGACCGCTGGGCAATGGAAATTTGGCTCCAAACAGACGAC 2280
Qy 2281 CTCGCCGATGTCAACAACAGATTCGGGAGCTGTCAATACAGTGCAGAGGGGAGCGCTCC 2340
Db 2281 CTCGCCGATGTCAACAACAGATTCGGGAGCTGTCAATACAGTGCAGAGGGGAGCGCTCC 2340
Qy 2341 CTGAGAAATGTGACCACTTACCTAATATCTTGGGAGTTTAAACAAGAACTCACAGGC 2400
Db 2341 CTGAGAAATGTGACCACTTACCTAATATCTTGGGAGTTTAAACAAGAACTCACAGGC 2400
Qy 2401 TAGCCGTGGGCAAGGGGTATGTCGCAATGACAGAAATGCAATCTTCAATCAAGTGTCCC 2460
Db 2401 TAGCCGTGGGCAAGGGGTATGTCGCAATGACAGAAATGCAATCTTCAATCAAGTGTCCC 2460
Qy 2461 GGAGCAAGGCTGAGGATCTTATGAGCAAGTCCAAACTGATGCGAAGATGAAGAAT 2520
Db 2461 GGAGCAAGGCTGAGGATCTTATGAGCAAGTCCAAACTGATGCGAAGATGAAGAAT 2520
Qy 2521 GATCATAGAGTAATTTCCATGAAGACTGAGAGTCAATCAAGTGTGATCGAGGACAG 2580
Db 2521 GATCATAGAGTAATTTCCATGAAGACTGAGAGTCAATCAAGTGTGATCGAGGACAG 2580
Qy 2581 GATTTATGTGACTACTGACAGATTCGAATCTGTATTTGACAGCAATTTGTGACAT 2640
Db 2581 GATTTATGTGACTACTGACAGATTCGAATCTGTATTTGACAGCAATTTGTGACAT 2640
Qy 2641 CTCGCGAATGCTTGGACGTCCTGATAGAGAGTGCCAGAGTCCGTGCAACTCGTG 2700
Db 2641 CTCGCGAATGCTTGGACGTCCTGATAGAGAGTGCCAGAGTCCGTGCAACTCGTG 2700
Qy 2701 GACTTCTGAAACCCCACTATCATGCGGCAAGGTCTCTGGGAAACCCAGAGTCCCA 2760
Db 2701 GACTTCTGAAACCCCACTATCATGCGGCAAGGTCTCTGGGAAACCCAGAGTCCCA 2760
Qy 2761 GTGACAGAGGCCAGCTTTGTGTAACTGCGTTTGAACCTTGCGGAGAACTCCCAAG 2820
Db 2761 GTGACAGAGGCCAGCTTTGTGTAACTGCGTTTGAACCTTGCGGAGAACTCCCAAG 2820
Qy 2821 CTAGCCAGGCTGAGGCTTCAAGCCAGCTACCGAGAGAGCAATGGCGAGCTGTGGGG 2880
Db 2821 CTAGCCAGGCTGAGGCTTCAAGCCAGCTACCGAGAGAGCAATGGCGAGCTGTGGGG 2880
Qy 2881 TCAGGCGGTATAGACAGCAGAGAGACTTCTGTGTGTGTCAGGCTTCTTCCAGAAC 2940
Db 2881 TCAGGCGGTATAGACAGCAGAGAGACTTCTGTGTGTGTCAGGCTTCTTCCAGAAC 2940
Qy 2941 ATTCAGACTCCCTGTCTGTGGGGAATGGGCTCCAGATACGTCCTTCTTGGCCAGACTGC 3000
Db 2941 ATTCAGACTCCCTGTCTGTGGGGAATGGGCTCCAGATACGTCCTTCTTGGCCAGACTGC 3000
Qy 3001 ATTCACCAATATGAGAAATTCATCTCCAGCTGCGAGAGCCCTTGGAGCAATATGCTT 3060
Db 3001 ATTCACCAATATGAGAAATTCATCTCCAGCTGCGAGAGCCCTTGGAGCAATATGCTT 3060
Qy 3061 GAAACAATTGAGAGCAAAACAGAGACCTTGAAGCTTGAAGACAGAGATGCCATCACTGT 3120
Db 3061 GAAACAATTGAGAGCAAAACAGAGACCTTGAAGCTTGAAGACAGAGATGCCATCACTGT 3120
Db 3061 GAAACAATTGAGAGCAAAACAGAGACCTTGAAGCTTGAAGACAGAGATGCCATCACTGT 3120
Qy 3121 CCACACTAGAAATGAGCCCTTCCGAGAACCCCTCGGAATGATACATCACTGATCCCATC 3180
Db 3121 CCACACTAGAAATGAGCCCTTCCGAGAACCCCTCGGAATGATACATCACTGATCCCATC 3180
Qy 3181 AAGCCAGCCATTGAGAACTGGGGCAGTGAATTCCTGTGTACAGAGTGAAGGCTTCAAT 3240
Db 3181 AAGCCAGCCATTGAGAACTGGGGCAGTGAATTCCTGTGTACAGAGTGAAGGCTTCAAT 3240
Qy 3241 AGTGTTCAACTCTGTCCAACAGCTCCGACAGACGACATCAAGTGTGCGCCCTTG 3300
Db 3241 AGTGTTCAACTCTGTCCAACAGCTCCGACAGACGACATCAAGTGTGCGCCCTTG 3300
Qy 3301 GGTGACTCTGATACAGAGTGGGAGTGCAGCCAAACATCCAGTGAACCTTACACCA 3360
Db 3301 GGTGACTCTGATACAGAGTGGGAGTGCAGCCAAACATCCAGTGAACCTTACACCA 3360
Qy 3361 TCTTGGAGGGGACTCTCTTGGAGCAATGGAGGGGATGGGAACTTGGAGACTGACACACA 3420
Db 3361 TCTTGGAGGGGACTCTCTTGGAGCAATGGAGGGGATGGGAACTTGGAGACTGACACACA 3420
Qy 3421 CTGCCCAACATTTGAGAAATTTCAACCTTACCTTCTTGGCTTCTTACAGACACTGG 3480
Db 3421 CTGCCCAACATTTGAGAAATTTCAACCTTACCTTCTTGGCTTCTTACAGACACTGG 3480
Qy 3481 GAGGGGACAGAGACTTAAATGGCAAGGGGAGGGGCGCAGAGCTTGGGACATGCCAGCC 3540
Db 3481 GAGGGGACAGAGACTTAAATGGCAAGGGGAGGGGCGCAGAGCTTGGGACATGCCAGCC 3540
Qy 3541 CAGGCTGGGACCTGTGAGAGGAAATGAAGAAACAGCCCGACATCACTTGAAGAAAGAC 3600
Db 3541 CAGGCTGGGACCTGTGAGAGGAAATGAAGAAACAGCCCGACATCACTTGAAGAAAGAC 3600
Qy 3601 TGAAGCTGTGACACTTTCATTTGGGGTCAACGACTTGTGATTACTGTGAGAAATCCG 3660
Db 3601 TGAAGCTGTGACACTTTCATTTGGGGTCAACGACTTGTGATTACTGTGAGAAATCCG 3660
Qy 3661 GAGGCCCAATTTGGGCAAGGAAATGTTGACGACATTCGCAAGCCCTTGACATCTCTCT 3720
Db 3661 GAGGCCCAATTTGGGCAAGGAAATGTTGACGACATTCGCAAGCCCTTGACATCTCTCT 3720
Qy 3721 GAGGAGCTCCCAAGGCTTTCGTCAACGTTGTGAGAGTCAATGAGAGCTGTGACCTGTAC 3780
Db 3721 GAGGAGCTCCCAAGGCTTTCGTCAACGTTGTGAGAGTCAATGAGAGCTGTGACCTGTAC 3780
Qy 3781 CAGGGCCTAAGGCGGGAATGTGCAATGCTGAGAGCTCAGAACATCTGCTCCAG 3840
Db 3781 CAGGGCCTAAGGCGGGAATGTGCAATGCTGAGAGCTCAGAACATCTGCTCCAG 3840
Qy 3841 CACTGCGAAAGCTCCCTGAGAGAGCAAGATGGAAGAAATGTAATGGAACCTCCAGCAT 3900
Db 3841 CACTGCGAAAGCTCCCTGAGAGAGCAAGATGGAAGAAATGTAATGGAACCTCCAGCAT 3900
Qy 3901 GGCATCTCCAGTTTCTCTCACTGAGCAACATCACAGCGTGAAGGACTTTCGGGTTGG 3960
Db 3901 GGCATCTCCAGTTTCTCTCACTGAGCAACATCACAGCGTGAAGGACTTTCGGGTTGG 3960
Qy 3961 GTGACAGCTTCTTCCAAACACATCAACCCCACTGAAGAGAGAGGGGACACTGACCTC 4020
Db 3961 GTGACAGCTTCTTCCAAACACATCAACCCCACTGAAGAGAGAGGGGACACTGACCTC 4020
Qy 4021 ACCTTCTTCTCGAGAGAGCTTTTCACTTCTAGACCGGGGCAATGCCAGATGCCATC 4080
Db 4021 ACCTTCTTCTCGAGAGAGCTTTTCACTTCTAGACCGGGGCAATGCCAGATGCCATC 4080
Qy 4081 GCACTCTGGAACAATGCTGGAACCAAGTGGGCGGCAAGACTTACCTCAACAACTTCAAC 4140
Db 4081 GCACTCTGGAACAATGCTGGAACCAAGTGGGCGGCAAGACTTACCTCAACAACTTCAAC 4140
Qy 4141 CACAGCCGAGCCAACTCAAGTGCCTCTCTCTGAGAGCCCTTACCTTACACCTGTGCGG 4200
Db 4141 CACAGCCGAGCCAACTCAAGTGCCTCTCTCTGAGAGCCCTTACCTTACACCTGTGCGG 4200
```

QY 4201 AACAGCCGATTTGCTCCAGACCAAGGCTGAAGAGCCCGAGGNGCTCTACTGAGGCTGTC 4260
DB 4201 AACAGCCGATTTGCTCCAGACCAAGGCTGAAGAGCCCGAGGNGCTCTACTGAGGCTGTC 4260
QY 4261 CCAAGTGGACGGGAGTGGGCTTTGTTGGTGGGATCATCGGAGACAGTGTCTGAGAGTGC 4320
DB 4261 CCAAGTGGACGGGAGTGGGCTTTGTTGGTGGGATCATCGGAGACAGTGTCTGAGAGTGC 4320
QY 4321 AGAGAGAGTGGCCGGAGGAGAGATCTCTCAATGAGCTTGGGACATGTGGCCCTCTAG 4377
DB 4321 AGAGAGAGTGGCCGGAGGAGAGATCTCTCAATGAGCTTGGGACATGTGGCCCTCTAG 4377
RESULT 2
AX417817 3648 bp DNA linear PAT 18-JUN-2002
LOCUS Sequence 1 from Patent WO0231161.
AX417817
VERSION AX417817.1 GI:21522937
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Zhu, Z.
TITLE Regulation of human phospholipase-like enzyme
JOURNAL Patent WO 0231161-A 1 18-Apr 2002;
BAYER AG (DE)
FEATURES
source 1..3648
location/qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 899 a 1032 c 947 g 770 t
ORIGIN
Query Match 79.1%; Score 3461.2; DB 6; Length 3648;
Best Local Similarity 97.5%; Pred. No. 0;
Matches 3558; Conservative 0; Mismatches 33; Indels 57; Gaps 2;
QY 775 GAAGCCGGAACAGCCTCTGCTGCTCCAGAGGATGAGAGAGAGTCTTCAACGCTG 834
DB 1 GAAGCCGGAACAGCCTCTGCTGCTCCAGAGGATGAGAGAGAGTCTTCAACGCTG 60
QY 835 GTTTCAGGCTTTCTTCTATGAGACCAACCCCATCTGTACACTGAGAGACCCCGAATC 894
DB 61 GTTTCAGGCTTTCTTCTATGAGACCAACCCCATCTGTACACTGAGAGACCCCGAATC 120
QY 895 CAGGATTTTACACGCTGGGCTGCACTCTGGAATGAGATGAGAGCCAGAGAGAG 954
DB 121 CAGGATTTTACACGCTGGGCTGCACTCTGGAATGAGATGAGAGCCAGAGAGAG 180
QY 955 AAAGATGAGCATTTGAGTGTAAAAACAGGAGGCAATGAGTGTCCCTCTCAGAGAGC 1014
DB 181 AAAGATGAGCATTTGAGTGTAAAAACAGGAGGCAATGAGTGTCCCTCTCAGAGAGC 240
QY 1015 CCCTATCTGTACGCTACAGAAACAGAACTACCTGACCAAGCTGAGAAACCCCAAGAC 1074
DB 241 CCCTATCTGTACGCTACAGAAACAGAACTACCTGACCAAGCTGAGAAACCCCAAGAC 300
QY 1075 AAGCTTG-----AGTTAAGAA 1092
DB 301 AAGCTTGAGAAAAATCATTTTCTTATTTCCCTGTGTCTCATTTCAAGGTAAAGAA 360
QY 1093 GAGAGCGAATTCAGATGCTCCGACAAAGACCCCTCGATAGGTTCCACCTGAGTTCA 1152
DB 361 GAGAGCGAATTCAGATGCTCCGACAAAGACCCCTCGATAGGTTCCACCTGAGTTCA 420
QY 1153 AGGCTGAGAGCGGCTGACATCAAGTAATTGAGAGCCCTGGGTGACTCTTCAAGGAGG 1212
DB 421 AGGCTGAGAGCGGCTGACATCAAGTAATTGAGAGCCCTGGGTGACTCTTCAAGGAGG 480

QY 1213 AATGGGCGGGGTCACACCTGGGAAGCTTTGAGAGCTTGAATGATGAGAGGCTG 1272
DB 481 AATGGGCGGGGTCACACCTGGGAAGCTTTGAGAGCTTGAATGATGAGAGGCTG 540
QY 1273 TCCTGAGGCTGGCGGAGATGAGAAATCATCGGACCGTTATCCACCTGGGAAATCTTC 1332
DB 541 TCCTGAGGCTGGCGGAGATGAGAAATCATCGGACCGTTATCCACCTGGGAAATCTTC 600
QY 1333 CGGGAATTCAACCTTCCCTGAAAGGGCTTCTGTGTGAGCATGGGAAAGAACCAATCTT 1392
DB 601 CGGGAATTCAACCTTCCCTGAAAGGGCTTCTGTGTGAGCATGGGAAAGAACCAATCTT 660
QY 1393 AATGCTTCTTAACAGGCTGTGGAGAGGCGGAGCTGAGGATCTACCTGTCCAGGCT 1452
DB 661 AATGCTTCTTAAACAGGCTGTGGAGAGGCGGAGCTGAGGATCTACCTGTCCAGGCT 720
QY 1453 AGGAGCTGGTGAACCTGATGAGAAATGACACGAGATACACTTTCAGGAAAGCTGAA 1512
DB 721 AGGAGCTGGTGAACCTGATGAGAAATGACACGAGATACACTTTCAGGAAAGCTGAA 780
QY 1513 ATATAAACCCTGTTATAGGCGGCAATGACCTGTGATTTCTGCAATGATCTGTCCAC 1572
DB 781 ATATAAACCCTGTTATAGGCGGCAATGACCTGTGATTTCTGCAATGATCTGTCCAC 840
QY 1573 TATTCTCCAGAACTTCAAGACAACTTGAAGAGGCGCTGACATCTCCATGCTGAG 1632
DB 841 TATTCTCCAGAACTTCAAGACAACTTGAAGAGGCGCTGACATCTCCATGCTGAG 900
QY 1633 GTTCTCGGCAATTTGTGAACCTGTGACGCTGTTGAGATCTGCAACCTGAGGAGCTG 1692
DB 901 GTTCTCGGCAATTTGTGAACCTGTGACGCTGTTGAGATCTGCAACCTGAGGAGCTG 960
QY 1693 TACCAGAGAAAAAAGTCTACTGCGGCAAGATATCTCAGGCTCTGTGTCTCTGTGTC 1752
DB 961 TACCAGAGAAAAAAGTCTACTGCGGCAAGATATCTCAGGCTCTGTGTCTCTGTGTC 1020
QY 1753 CTGAAGTTGATGATTAACCAAGAACTTGTCAACCTCATGCAATTCACAAAGAGTTT 1812
DB 1021 CTGAAGTTGATGATTAACCAAGAACTTGTCAACCTCATGCAATTCACAAAGAGTTT 1080
QY 1813 CAGAGAAAGCCACCAACTGATTTGAGAGTGGGCGATATGACAAAGGAAATTTTACT 1872
DB 1081 CAGAGAAAGCCACCAACTGATTTGAGAGTGGGCGATATGACAAAGGAAATTTTACT 1140
QY 1873 GTGCTGTGACGCTTTCTTTGAAAACGTGACATGCAAGACCTTGGAAAGATTGCT 1932
DB 1141 GTGCTGTGACGCTTTCTTTGAAAACGTGACATGCAAGACCTTGGAAAGATTGCT 1200
QY 1933 GACAACTTTCTTGGCTCTGCTGCTGCTTTTCACTTCAAGAGCAAGTCTACCTCCGAGCA 1992
DB 1201 GACAACTTTCTTGGCTCTGCTGCTGCTTTTCACTTCAAGAGCAAGTCTACCTCCGAGCA 1260
QY 1993 GCCAGTCTCTCTGAAACATATGCTGAGCCTGTGTGGCCAGAAAGCACTGCTATAG 2052
DB 1261 GCCAGTCTCTCTGAAACATATGCTGAGCCTGTGTGGCCAGAAAGCACTGCTATAG 1320
QY 2053 TTTGAAAACAGATCAATATCATATGCTGAGACAGTCCAGGCTTTCTGAGAGCTTAC 2112
DB 1321 TTTGAAAACAGATCAATATCATATGCTGAGACAGTCCAGGCTTTCTGAGAGCTTAC 1380
QY 2113 AAGAACG-----CATGCAAGGCTGAGGAGCCTGGGCTGAGGAGCTGAGGAG 2157
DB 1381 AAGAACG-----CATGCAAGGCTGAGGAGCCTGGGCTGAGGAGCTGAGGAG 1440
QY 2158 AGAGCCCTTCTGCTTGGACCTTACCTGATGAGTATGAGTATGAGTATGAGTATGAG 2217
DB 1441 AGAGCCCTTCTGCTTGGACCTTACCTGATGAGTATGAGTATGAGTATGAGTATGAG 1500
QY 2218 GTTGTGGCTGCTGTGGGGAATTTCTGACCGCTGAGCAATGAAATGGCTTCAACAGAC 2277
DB 1501 GTTGTGGCTGCTGTGGGGAATTTCTGACCGCTGAGCAATGAAATGGCTTCAACAGAC 1560

QY 2278 GACCTCCCGGAGTGCACACAGATCGGGAGCTGTATACATGACGAGAGGGGAGCGC 2237
DB 1561 GACCTCCCGGAGTGCACACAGATCGGGAGCTGTATACATGACGAGAGGGGAGCGCG 1620
QY 2338 TCCCTGGAGATGAGACCACTTAATATCTCTGGGAGTTTAAACAAAACCTCAC 2297
DB 1621 TCCCTGGAGATGAGACCACTTAATATCTCTGGGAGTTTAAACAAAACCTCAC 1680
QY 2398 GGGTACGCGTGGGACGCGGTGATGACATGACAGATGATCTCTCAATCAAGCTGT 2457
DB 1681 GGGTACGCGTGGGACGCGGTGATGACATGACAGATGATCTCTCAATCAAGCTGT 1740
QY 2458 CCGGAGCAAAAGGCTGAGATCTTATGAGCAAGTCCAACTCTGATGACAGAGTAAA 2517
DB 1741 CCGGAGCAAAAGGCTGAGATCTTATGAGCAAGTCCAACTCTGATGACAGAGTAAA 1800
QY 2518 GATGATCATGATGATTAATTTCCATGAAAGCTGAAAGCTCATCACAGTGTGATGAGG 2577
DB 1801 GATGATCATGATGATTAATTTCCATGAAAGCTGAAAGCTCATCACAGTGTGATGAGG 1860
QY 2578 AGCGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2637
DB 1861 AGCGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
QY 2638 CATCTCCGCAATGCTTGGAGCTCTGATGAGAGAGGCTCCAGAGTCTGTCAACCTC 2697
DB 1921 CATCTCCGCAATGCTTGGAGCTCTGATGAGAGAGGCTCCAGAGTCTGTCAACCTC 1980
QY 2698 GTGGACTTCTGAAACCCCACTATATGCGGAGAGTGTCTCTGGGAAAACCCAGCAAGTGC 2757
DB 1981 GTGGACTTCTGAAACCCCACTATATGCGGAGAGTGTCTCTGGGAAAACCCAGCAAGTGC 2040
QY 2758 CCGATGAGAGGCGAGCGTGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2817
DB 2041 CCGATGAGAGGCGAGCGTGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 2100
QY 2818 GAGCTAGCAGGCTGAGAGGCTTCAAGCCAGCTTACCGAGAGCAATGCGAGCTGTG 2877
DB 2101 GAGCTAGCAGGCTGAGAGGCTTCAAGCCAGCTTACCGAGAGCAATGCGAGCTGTG 2160
QY 2878 GGGTCAAGGCGCTATGACACGAGAGAGATTTCTGTGTGTGCTGAGCCCTTCTTCAG 2937
DB 2161 GGGTCAAGGCGCTATGACACGAGAGAGATTTCTGTGTGTGCTGAGCCCTTCTTCAG 2220
QY 2938 AACATCAGCTCCCTCTCTGCGGAGTGGGCTCCAGATACGCTCTTCTTGGCCCAAGAC 2997
DB 2221 AACATCAGCTCCCTCTCTGCGGAGTGGGCTCCAGATACGCTCTTCTTGGCCCAAGAC 2280
QY 2998 TGCATCCACCAATCAGAAATTCATCTCCAGCTGGCCAGAGCCCTTGGACCAATATG 3057
DB 2281 TGCATCCACCAATCAGAAATTCATCTCCAGCTGGCCAGAGCCCTTGGACCAATATG 2340
QY 3058 CTTGAAACCTTGAAAGCAAAACAGAGACCTTGAGCTGAGAGCAGAGATGCCATCAC 3117
DB 2341 CTTGAAACCTTGAAAGCAAAACAGAGACCTTGAGCTGAGAGATGCCATCAC 2400
QY 3118 TGTCCCACTCAGAAATGAGCCCTTCTGAGAAACCTTGGAAATGTAATCAACATACCCC 3177
DB 2401 TGTCCCACTCAGAAATGAGCCCTTCTGAGAAACCTTGGAAATGTAATCAACATACCCC 2460
QY 3178 ATCAAGCCAGCATGAGAACTGGGCAAGTGAATCTCTGTGTGATCAGAGTGAAGGCTTC 3237
DB 2461 ATCAAGCCAGCATGAGAACTGGGCAAGTGAATCTCTGTGTGATCAGAGTGAAGGCTTC 2520
QY 3238 AATAGTGTTCAACTCTGTGTCACAGCTCCGACAGACATCAAAAGTGTGAGCGGC 3297
DB 2521 AATAGTGTTCAACTCTGTGTCACAGCTCCGACAGACATCAAAAGTGTGAGCGGC 2580
QY 3298 CTGGGAGACTCTGTGATCAGAGAGTGGAGCTGACCAAACTCTCACTGATCACTACC 3357
DB 2581 CTGGGAGACTCTGTGATCAGAGAGTGGAGCTGACCAAACTCTCACTGATCACTACC 2640
QY 3358 ACATCTTGAAGGAGACTCTGTGAGAGATGAGAGGAGTGAAGTCTTGAAGTCACTAC 3417

DB 2641 ACATCTTGAAGGAGACTCTGTGAGAGATGAGAGGAGTGAAGTCTTGAAGTCACTAC 2700
QY 3418 ACATCTTGAAGGAGACTCTGTGAGAGATGAGAGGAGTGAAGTCTTGAAGTCACTAC 3477
DB 2701 ACATCTTGAAGGAGACTCTGTGAGAGATGAGAGGAGTGAAGTCTTGAAGTCACTAC 2760
QY 3478 TGGAGAGGAGACAGAGAGTAAATGTGGACGGAAGGGGCGAGCTGAGGACATGCCA 3537
DB 2761 TGGAGAGGAGACAGAGAGTAAATGTGGACGGAAGGGGCGAGCTGAGGACATGCCA 2820
QY 3538 GCCAGGCTGGAGCTGTGAGAGCAATGAAAGAGCCCGGACATCACTGAGAGAA 3597
DB 2821 GCCAGGCTGGAGCTGTGAGAGCAATGAAAGAGCCCGGACATCACTGAGAGAA 2880
QY 3598 GACTGAGAGTGTGACACTCTTCAATGGGAGTAAAGACTTGTGTATTAAGTGAAT 3657
DB 2881 GACTGAGAGTGTGACACTCTTCAATGGGAGTAAAGACTTGTGTATTAAGTGAAT 2940
QY 3658 CCGAGAGCCCACTTGGACAGGAATATGTTCAAGCATCAACAGGCTGGACATCTC 3717
DB 2941 CCGAGAGCCCACTTGGACAGGAATATGTTCAAGCATCAACAGGCTGGACATCTC 3000
QY 3718 TGTGAGAGCTCCCAAGGCTTCTGTAACATGTGTGAGAGTCAATGAGCTGCTAGCTC 3777
DB 3001 TGTGAGAGCTCCCAAGGCTTCTGTAACATGTGTGAGAGTCAATGAGCTGCTAGCTC 3060
QY 3778 TACAGAGGCGAAGGCGGAAATGTGCAATGCTGAGAGCTGCAAGCACTGCTC 3837
DB 3061 TACAGAGGCGAAGGCGGAAATGTGCAATGCTGAGAGCTGCAAGCACTGCTC 3120
QY 3838 AGACACTGCAAAAGCTCCCTGAGAGAGCAAGATGAAAGTGAAGTGAAGCTCCAG 3897
DB 3121 AGACACTGCAAAAGCTCCCTGAGAGAGCAAGATGAAAGTGAAGTGAAGCTCCAG 3180
QY 3898 CATGCACTCCAGTCTTCTCTACTGAGCAACATACACAGAGTGAAGCTTGGGTT 3957
DB 3181 CATGCACTCCAGTCTTCTCTACTGAGCAACATACACAGAGTGAAGCTTGGGTT 3240
QY 3958 GTGTGAGAGCTTCTTCCAAACACATCAACCTGAAAGAGAGGAGCACTGAC 4017
DB 3241 GTGTGAGAGCTTCTTCCAAACACATCAACCTGAAAGAGAGGAGCACTGAC 3300
QY 4018 CTCACTTCTTCCAGAGAGCTTCTCTGAGAGAGGAGAGTCCAGATGAGCC 4077
DB 3301 CTCACTTCTTCCAGAGAGCTTCTCTGAGAGAGGAGAGTCCAGATGAGCC 3360
QY 4078 ATGCACTCTGGAACCAATGCTGGAACAGTGGGCGGCAAGCTCTCAACACTTC 4137
DB 3361 ATGCACTCTGGAACCAATGCTGGAACAGTGGGCGGCAAGCTCTCAACACTTC 3420
QY 4138 ACCCAGAGCGAGCCAACTCAAGTGCCTTCTGAGAGCCCTTACCTCAACCTG 4197
DB 3421 ACCCAGAGCGAGCCAACTCAAGTGCCTTCTGAGAGCCCTTACCTCAACCTG 3480
QY 4198 CGGAAGAGCGATGTCTCCCAAGACAGAGTGAAGAGCCCGAGAGTCTTACTGGGCT 4257
DB 3481 CGGAAGAGCGATGTCTCCCAAGACAGAGTGAAGAGCCCGAGAGTCTTACTGGGCT 3540
QY 4258 GTCCAGTGGCAGAGGAGTGGGCTTGTGTGAGGATCATCGGAGCAATGATCTGAGG 4317
DB 3541 GTCCAGTGGCAGAGGAGTGGGCTTGTGTGAGGATCATCGGAGCAATGATCTGAGG 3600
QY 4318 TGCAGAGAGTGGCTCGAGAGGAATCTCTCAATGAGCTGCGCACT 4365
DB 3601 TGCAGAGAGTGGCTCGAGAGGAATCTCTCAATGAGCTGCGCACT 3648

RESULT 3
OCPPHLP 4612 bp mRNA linear MAM 09-JUN-1994
LOCUS OCPPHLP
DEFINITION O.cuniculus mRNA for phospholipase.
ACCESSION Z12841

VERSION 212841.1 GI:1689
 KEYWORDS phosholipase.
 SOURCE Oryctolagus cuniculus (rabbit)
 ORGANISM Oryctolagus cuniculus
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 REFERENCE 1
 Bolli, W., Schmid-Chanda, T., Semenza, G. and Manzel, N.
 Messenger RNAs expressed in intestine of adult but not baby
 rabbits. Isolation of cognate cDNAs and characterization of a novel
 brush border protein with esterase and phospholipase activity
 J. Biol. Chem. 268 (17), 12901-12911 (1993)
 MEDLINE 93286138
 PUBMED 8509424
 REFERENCE 2 (bases 1 to 4612)
 Bolli, W.
 Direct Submission
 Submitted (25-JUN-1992) WERNER BOLLI, Dept. of Biochemistry 11,
 Swiss Federal Institute of Technology, ETH Zentrum,
 Universitaetstrasse 16, Zuerich, CH-8092, Switzerland
 Location/Qualifiers
 1..4612
 /organism="Oryctolagus cuniculus"
 /mol_type="mRNA"
 /strain="New Zealand White"
 /db_xref="taxon:9986"
 /clone="padrab-1113"
 /cell_type="Epithelial cell"
 /tissue_type="Small intestine"
 /dev_stage="Adult"
 25..4401
 /codon_start=1
 /product="Phospholipase"
 /protein_id="CAA78303.1"
 /db_xref="GI:1690"
 /translation="MALPSPVFLGLPLPLRGADQIOTSSGNKLTLEGOLMPSLKT
 PPGDPTLAEVSPESVSLRPDIKEVAIGVETPAVAGADIMQATLEEDQETKRPD
 ACHGVTVLSDIIGRESPSALMPCETPTLPPRGADLMQATELVRSKRENPQDF
 EHMVKLVNFFSNTSCPCPSAQKLVLAGMDKTRTLDYVLAQVPAFVNLVDLS
 ELAFAFRMQAOLSPAEPCRCLETSOLTKVLTOMSVLEANDSLASRYNDSF
 AVVQEPFESSLSALAEPLDPTLALSLMNMMEPIKREKPESEKERYLRCP
 TQESPLYFTYRNSGLTRVPOGKLEVRGTEICPDQSDSPVTRHRLPADIK
 VIGAMGSLTAGNGASGSPGNILDLTVQVGLSVSGGDSDNISTVTLAILLBPNS
 LQFSVCTGRETTSQAFNPAVAGARADGILPQARVVALMKDTRINPEBMTITV
 FIVGDLCDPCNDVRSYFONFTDNISTADILHAEIPRAVNLVKLETSLKRELQ
 ETQVSCPRMILRSICPCVLKFDNSTEIASLIEIKYKORRTOQLIDSGRYDRDPT
 VVLAQPEKVMKPTQDGLPDSNFADCFHFSKAHAHAASALMNMLEPVQKXTH
 NDEGAVNITCPNQMWPFLSTYKNSVQGFGLWPCRDSPSASPPTSVHARPADIV
 VALAGDSLTAIGTIGSKPNDLSGTOYRGLSYSSGGDSLNVTTTPIILROPNSL
 MGRAVGTGDSGTRAFPNQAVPGAKADLMSQVOTLVQRMKDRHNVFQEDMKITVQ
 IGASDLDCYCTDSNLYSANFVYDHLRALDLHREVRALVNLDFNVPDVKOVFG
 NPKPCPVQOASALCNCLVSPRENSYELARLEALAOVSSLRLEVESGRTVEDFV
 VLQPFPHSIOLPVLQDGLDTSFPAPCVHFNOKFHSOLSRALRMNLEPLGGKTDL
 DLQALITLCPTONEPLRTFRNSDYTPSPAVENMGSDPLCTANNAASRGVSVE
 LQGGDITVVALADSLTLAMGARSNSDDPMFRGSSNISGGGALERTTTPNLTK
 KFNPSILIGSTGIEGTGLAVAVQGARADMPAQAADLVKRMNSPELIDKWKV
 TLVPGNDLCHFCENPBGSSBGEVYOHIOALDVLIELPRTFVNVVEVNEGLHAD
 QGGKCATLLAOSHCTCFKYSOSVEMQELKVMNLSQGLSLSYSHQVVOGEDFV
 VVOFPONTLVPLNRGDTDLTFPSDDCFHSGHAEMIALMNMNLEPVGHXTTSM
 NPTYSRTKLCPSPDSPLYLTILRSRLLPQAEADPTVLYMAVPAAGAGLLIGILAM
 VAGRMRCRREDPPLSLSTGL"
 25..81
 sig_peptide
 polyA_site
 BASE COUNT 1031 a 1414 c 1293 g 874 t
 ORIGIN
 Query Match 71.4%; Score 3124; DB 4; Length 4612;
 Best Local Similarity 82.6%; Pred. No. 0;
 Matches 3613; Conservative 0; Mismatches 750; Indels 9; Gaps 3;
 1 ATGGGCTGGGCGGCAAGCATTTCTCTCTGAGAGTGTCTGCTTCTGGGCGAAGGACC 60

Db 25 ATGGGCTGGGCGGCAAGCATTTCTCTCTGAGAGTGTCTGCTTCTGGGCGAAGGACC 84
 Qy 61 CCTCAAGATCATACCTCTCTCTAGAAAAGATCACTTGAAGAGGAGCTATGAGCCAGACC 120
 Db 85 GATCAGATCCAACTCTCTCTGAGAAACACATTTGAAAGGGAGCTCTGGCCAGAGAGC 144
 Qy 121 CTGAAGATCTCTCATTTCCATGCAACCCAAATTAATTTGAGTGAATATGCTTCTTAA 180
 Db 145 CTGAAGATCTTCTCATTTCCATGCAACCCAAATTAATTTGAGTGAATATGCTTCTTAA 204
 Qy 181 TCAGTTCATCTCTGAAAGCTTCTGATATTAATTTGAGCAGCATTTGCAATCTGAA 240
 Db 205 TCAGTTCATCTCTGAAAGCTTCTGATATTAATTTGAGCAGCATTTGCAATCTGAA 264
 Qy 241 ATTCCTCAGACCCAGGAGGAGGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 297
 Db 265 ACTGCTCCAGACTCAGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 324
 Qy 298 CAGCAGGTCGATGGAG 357
 Db 325 GAGCAGGTCGATGGAG 384
 Qy 358 TCTGTTCCATGCTGCTGTGTGTCACACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 417
 Db 385 TCTGTTCCATGCTGCTGTGTGTCACACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 441
 Qy 418 GACTTGTGATTCAGGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 477
 Db 442 GACTTGTGATTCAGGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 501
 Qy 478 TTTCAATTGACATGGAAGCTCATCAATGTGTTCTTCAATGATGCAAGCCAGTTCCTG 537
 Db 502 TTTGAACATGACATGGAAGCTCATCAATGTGTTCTTCAATGATGCAAGCCAGTTCCTG 561
 Qy 538 TGGCCCTCTGCTCAAGAGATGGGCTTGGCGGCGGCGGCTGTGATGAGCTGATGGGGGTG 597
 Db 562 TGGCCCTCTGCTCAAGAGATGGGCTTGGCGGCGGCGGCTGTGATGAGCTGATGGGGGTG 621
 Qy 598 CTGAGCTACTGCAAG 657
 Db 622 CTGAGCTACTGCAAG 681
 Qy 658 GTTGAAGAGTCTCTGTGATGATCAAGGACTTGGCTGAGCCCTGACCAAGAGCCCTGT 717
 Db 682 GTTGAAGAGTCTCTGTGATGATCAAGGACTTGGCTGAGCCCTGACCAAGAGCCCTGT 741
 Qy 718 AATTGCTCAAG 777
 Db 742 AGGTGCTTAAG 801
 Qy 778 GCTTGAAG 837
 Db 802 GCTTGAAG 861
 Qy 838 TTTCAAGCTTCTTCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 897
 Db 862 TTTCAAGCTTCTTCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 921
 Qy 898 GATTTCACAGCTGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 957
 Db 922 GATTTCACAGCTGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 981
 Qy 958 GATGAGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017
 Db 982 GATGAGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1041
 Qy 1018 TATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
 Db 1042 TATCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1101
 Qy 1078 CTGAGGTAAG 1137

Db 1102 CTCGAGGTGAGAGAGGAGCGAAATCCGCTGTCCGACAGAGACCCTTCGACTCGGTT 1161
Qy 1138 CCCACCTGATTGATAGGCTGAGCGGCTGATGACATGAAATGAGCCCTGGGTGAC 1197
Db 1162 CCCACCTGATTGATAGGCTGAGCGGCTGATGACATGAAATGAGCTGCAATGAGGTGAC 1221
Qy 1198 TCTCTCAGCGCAGGCAATGAGGCGGAGTCCACACTGGGAAAGCTCTTGGAGCTTGA 1257
Db 1222 TCCCTCAGCGCAGGCAATGAGGCGGAGTCCACACTGGGAAAGCTCTTGGAGCTTGA 1281
Qy 1258 CAGTACCGAGGCTGTCTCTGAGCGTGTGCGAGATGAGAAATCGGCACTGATCCACC 1317
Db 1282 CAGTACCGAGGCTGTCTCTGAGCGTGTGCGAGATGAGAAATCGGCACTGATCCACC 1341
Qy 1318 CTGGGCAATCTCTCGGGAATTCACCCCTTCCCTGAGAGGCTTCTGTGGCACTGGG 1377
Db 1342 CTGGGCAATCTCTCGGGAATTCACCCCTTCCCTGAGAGGCTTCTGTGGCACTGGG 1401
Qy 1378 AAGAAACAGTCTTAATGCTTCTTAAACAGAGCTGTGAGAGGCGGAGCTGAGAT 1437
Db 1402 AGAGAGACCACTTCCAGGCTTCTTCAACAGGCACTGGCGGGCGCCAGCCGATGG 1461
Qy 1438 CTACTGTCCAGGCGCAGAGGCTGTGTGACCTGATGAAAGATGACAGAGATACCTTT 1497
Db 1462 TTAAATCCCAAGGCGGCTGTGTGACCTGATGAAAGATGACAGAGATACCTTT 1521
Qy 1498 CAGGAAGCTGGAAGATATTAACCTGTTATAGGCGGCAATGACCTCTGTGATTTCTGC 1557
Db 1522 CAGGAAGATTTGGAAGATTAATACCGTGTCTATGAGGCAACGACCTGTGTATTTCTGC 1581
Qy 1558 AATGATCTGGTTCACATCTCTCCAGAACTTCAAGACAACTTGGAAAGGCGCTGGAC 1617
Db 1582 AATGATCCGGGTGGCTACTCTCCCGAGAACTTCAAGCAACATCGGAGAGGCGCTGGAC 1641
Qy 1618 ATCTCTCATGCTGAGGTTCTCTCGGCAATTTGTGAACCTGTGTGACGCTGTGAGATCTC 1677
Db 1642 ATCTCTCAGCTGAGATTTCCCGGGCTTTGTGACCTGTGTGAAGGCTGTGAGATCTC 1701
Qy 1678 AACCTGAGGAGCTGTATCCAGAGAAAAAGTCTACTGCCAAGATGATCTCAGGCTT 1737
Db 1702 AAGCTGAGAGAGCTGTATCCAGAGAACTTAAGTGAAGTGCCTCCAGATGATCTCAGGCTT 1761
Qy 1738 CTGTGTCCCTGTGCTGCTGAAGTTGATGATACCAAGACATGTGCTACCTCATGAA 1797
Db 1762 CTGTGTCCCTGTGCTGCTGAAGTTGATGATACCAAGACATGTGCTACCTCATGAA 1821
Qy 1798 TTCAAAGAGAGTTTCAAGAGAACCAACCACTGATTGAGAGTGGGCGATATGACAA 1857
Db 1822 AGCATCAAGAGATATCAGAGAGAGACCCAGCACTGATTGACAGCGGGCGATACAGACG 1881
Qy 1858 AGGGAAGATTTTACTGTGTGTGACAGCGCTTCTTGAAGAGTGACATGCCAAGACC 1917
Db 1882 AGGGAAGACTTCAAGGTGTGTGTGACAGCGCTTCTTGAAGAGTGACATGCCAAGACC 1941
Qy 1918 TCGGAAGATGTGCTGACAACTCTTCTGCTGCTGACAGTGTTCACACTTCAAGACAG 1977
Db 1942 CAGGAGGAGCTACCCGACAACTCTTCTGCGCGCAGCTGTCTTCCACTTCAAGACAG 2001
Qy 1978 TCTCACTCCGAGCAGCAGTGTCTCTGGAACAATATGCTGAGAGCTGTGGCAGAG 2037
Db 2002 GCCCAGCGCAGCAGCAGTGTGCGCTCTGGAACAATGCTGGAACCTGTGGCAGAG 2061
Qy 2038 AGCATCTGTCAATAGTTTGAAGAAAGATCAATATCAATGCTGCGAAGCAGGTCCAGCG 2097
Db 2062 ACAACACACATATGTTTGAAGAGCAGTCAATTAAGTGTCCAAACAGGTCTGGCG 2121
Qy 2098 TTTCTAGAGCCTACAAAGACAGATGCAAGGTCAATGGGACCTGGCTGCAATGCAAGGAC 2157
Db 2122 TTTCTAGATCTTACAAAGACAGCTGTGAGGCTTTGGGACCTGGCTGCAATGCAAGGAC 2181
Qy 2158 AGAGCCCTTCTGCTTGAACCTTACCTGATGATGCTGAGAGCTGAGAGCTGAGACATCCAA 2217
Db 2182 AGATCCCGTCTGCTGCGCCCACTCACTGATGACCGCTGTGAGACCGCGAGATATCCAA 2241

Qy 2218 GTTGTGCTGCTCTGGGGATTTCTGACCGGCTGCAATGGAATGGCTCCAAACGAGAC 2277
Db 2242 GTTGTGCTGCTCTGGGAATTTCTTAAGCTCTGGGATTTGAATGGCTCCAAACGAAAT 2301
Qy 2278 GACCTCCCGATGTACCAACAGATTCGGGACTGTCAATACGTGACAGAGGGAGCGC 2337
Db 2302 GACCTCTCAGATGGGACCACTGATTCGGGGACTGTCAATATGTTCAAGAGGGAGCGC 2361
Qy 2338 TCCCTGGAAGATGTACCACTTACTTAATATCTTGGGAGTTTAAACGAAACTTCACA 2397
Db 2362 TCCCTGGAAGATGTACCACTTACTTAATATCTTGGGAGTTTAAACGAAACTTCATG 2421
Qy 2398 GGTACGCGGTGGGACGAGGTGATGCAATGACAGAAATGCAATTCCTCAATCAAGCTGT 2457
Db 2422 GGTTCGCGGTGGGACGAGGTGATGCAATGACAGAAATGCAATTCCTCAATCAAGCTGT 2481
Qy 2458 CCGGAGCAAAAGCTGAGATCTTATGAGCCAAATGCAACTCTGATGAGAGATGAA 2517
Db 2482 CCGGAGCAAAAGCTGAGATCTTATGAGCCAAATGCAACTCTGATGAGAGATGAA 2541
Qy 2518 GATGATCATAGATTAATTTCCATGAAGCTGGAAGTCAATCAAGTGTGATGAGAGC 2577
Db 2542 GACGACCATAGATTAATTTCCATGAAGATGGAAGTCAATCAAGTGTGATGAGAGC 2601
Qy 2578 AGCATTTATGTACATGACAGATGCAATGGAATCTGATTTGACAGCCAACTTGTGAC 2637
Db 2602 AGTACCTGTGTACATGACAGATGCAATGGAATCTGATTTGACAGCCAACTTGTGAC 2661
Qy 2638 CATCTCCGAGATGCTTGAAGCTCTGCAATGAGAGGTGCCAGAGTCTGTGCAACTC 2697
Db 2662 CATCTCCGAGATGCTTGAAGCTCTGCAATGAGAGGTGCCAGAGTCTGTGCAACTC 2721
Qy 2698 GTGACCTTCTGAACCCCACTATCATGCGCAGGTGTTCTTGGGAAACCCAGCAAGTGC 2757
Db 2722 GTGACCTTCTGAACCCCACTATCATGCGCAGGTGTTCTTGGGAAACCCAGCAAGTGC 2781
Qy 2758 CCAAGTACAGAGCGGCTTTGTGTATCTGCTGTGACCTGTGGGAGAACTCCCA 2817
Db 2782 CCAAGTACAGAGCGGCTTTGTGTATCTGCTGTGACCTGTGGGAGAACTCCCTAT 2841
Qy 2818 GAGCTAGCGAGCTGAGGCTTTCAGCGAGCTTACCGAGAGCAATGCGAGCTGTG 2877
Db 2842 GAGCTAGCGGCTGAGGCTTTCAGCGAGCGGAGGACTTCTCGTGTGCTGACGCTTCTTCAC 2901
Qy 2878 GAGTCAAGCGCTATGACAGAGAGGACTTCTGTGTGTGTGACGCTTCTTCAG 2937
Db 2902 GAGTCAAGCGCTATGACAGAGAGGACTTCTGTGTGTGTGACGCTTCTTCAC 2961
Qy 2938 AACATCAAGCTCTGTCTCTGCGGATGGGCTTCCAGATAGTCTTCTTGGCCGAGAC 2997
Db 2962 AACATCAAGCTCTGTCTCTGCGGATGGGCTTCCAGATAGTCTTCTTGGCCGAGAC 3021
Qy 2998 TGCATCACCAATCAGAAATTCATCTCCAGCTGGCAGAGCCCTTGGAGCAATATG 3057
Db 3022 TGCATCACCAATCAGAAATTCATCTCCAGCTGGCAGAGCCCTTGGAGCAATATG 3081
Qy 3058 CTTGAACCACTTGAAGCAAAACAGAGACCTTGAACCTGAGAGCAGAGATGCCATCAC 3117
Db 3082 CTGAGACCCCTGGAGAGGAGAAACGAGTGCATGAGACTTGAACCGGCAACCTCAC 3141
Qy 3118 TGTCCACTGAGATAGAGCCCTTCTGAGAACCCCTGGAATGATATCAACATGACCC 3177
Db 3142 TGTCCACTGAGATAGAGCCCTTCTGAGAACCCCTTCCGAAACGATGATCAACCTGACCC 3201
Qy 3178 ATCAAGCCAGCATGAGAACTGGGGCAGTGAATCTCTGTATCAGAGTGAAGGCTTCC 3237
Db 3202 AGCAGCCAGCCGTGAGAACTGGGGCAGTGAATCTCTGTATCAGAGTGAAGGCTTCC 3261
Qy 3238 AATAGTGTTCACACTTGTTCACAGCTTCGACAGCAGACGATCAAGTGTGCGCGC 3297
Db 3262 AGGAGTGTTCGAACTCAATTCATGAGCTTCAACGAGAGATATCAAGTGTGAGTGC 3321

Matches 3511; Conservative -; 0; Mismatches 780; Indels 21; Gaps 6;

QY 1 ATGGGCTGGGCGAGGCAATTTCTCTGAGAGTGGCTGCTTCTGGGCGAAGGAGC 60
 Db ATGGGCTGGGCGAGGCGTCTTACTTGTGGGCTGCTGCTG--CTGGGCTAAGGAGTC 96
 QY 61 CCTCAGATCCATACTCTCTAGAAAGATACATTGGAGGGGAGCTATGGCCAGAGACC 120
 Db ACCCAATTCACACATCTTCTGGAAAGACACTGGAGGCCAGCTATGCGCCAGAGACC 156
 QY 121 CTGAAGAAATCTCATTTCCATGCAACCCAAATTAATTAGAGTGAATAGCTTCTTAA 180
 Db AAAAAAATCTTCATCTCTGTAGAGCCAAAGAGCTGGAGCTGAACATGCTTCTGAG 216
 QY 181 TCAGTTCATCTCTGAGAGCTTCTGATATTAATTGTGGAGCAATGGCAATCTGAA 240
 Db TCAGTTCATCTCTGAGAGCTTCTGATATTAATTGTGGAGCAATGGCAATCTGAA 276
 QY 241 ATTCTTCAGACCCAGGAGCGGGGATCTGGAGAGCAAGCACTGAGCTG--AAAGGCA 297
 Db ACTCTTCAGACCTCTGGTGAAGTCACTGGAGCAAGGAGAGCTGAGAGAGAGCA 336
 QY 298 CAGCAGGTGTCATGGAGTGAATGACAGTCTTCAAGATCATCAGATATTTCAAGTCT 357
 Db TGCGGGGGGTGATGGGAATGATGACAGTCTTCAAGATCATCAGAGCTTCAAGCTT 396
 QY 358 TCGTTCATATGCTGTGTCCACATGGAAGAGATCATCCCGAGATGGTGTGAA 417
 Db TCGTTCATATGCTGTGTCCACATGGAAGAGCTGTGTGTGTGTGTGTGTGTGAA 456
 QY 418 GACTTGTGATGAGGCTCAAGAACTGTGAGAAACATGAAAGAGAGCTGCAACTGAC 477
 Db GAATTAAGGAGCTCAGGCTGAAGAGCTGTGAGAGCTTGAAGAGAGAGCCGCACTT 516
 QY 478 TTTCAATTGATGAGAACTCATATATGTGTCTTCAATATGCAAGCAAGTCTTAC 537
 Db TTTCAAGAGATTTGAAACTCATATATGTGTCTTCAATATGCAAGCAAGTCTTAC 576
 QY 538 TGCCCTCTGCTCAAGAAATGGGCTTGGCGGGGCGCTGATGAGTGAATGGGGGTG 597
 Db TGCCCTCTGCTCAAGAAATGGGCTTGGCGGGGCGCTGATGAGTGAATGGGGGTG 636
 QY 598 CTGAGCTACCTGAGAGAGAGTCTCCAGAGCAATTTGTAACCTGTGAGCTCTGAG 657
 Db CTACACTACCTGAGAGAGAGTCTCCAGAGCAATTTGTAACCTGTGAGCTCTGAG 696
 QY 658 GTTGCAGAGGTCTCTGTGATGATCAAGGCACTTGGCTGAG--CTGCAAGAGAGCC 714
 Db GTTGCAGAGGTCTCTGTGATGATGATCAAGGCACTTGGCTGAG--CTGCAAGAGAGCC 756
 QY 715 TGTAAATGCTCAAGAGAGACACCCGGCTGGCAAGGTGATGAGTGTCTTATCAG 774
 Db TGTAAATGCTCAAGAGAGACACCCGGCTGGCAAGGTGATGAGTGTCTTATCAG 816
 QY 775 GAAGCCTGGAAGAGCCTCTGGCTCCAGAGGAGTCAAGTGAAGAGAGTCTTCAAGC 834
 Db GAAGCCTGGAAGAGCCTCTGGCTCCAGAGGAGTCAAGTGAAGAGAGTCTTCAAGC 876
 QY 835 GTTTTCAGAGCTTCTTCTTATGAGAGCAACCTCTTACACTGAGAGAGAGCCCGA 894
 Db GTTTTCAGAGCTTCTTCTTATGAGAGCAACCTCTTACACTGAGAGAGAGCCCGA 930
 QY 895 CAGGATTTACACAGCTGGCTGGCACTCTGGAATGATGATGAGAGAGAGAGAG 954
 Db CAGGATTTACACAGCTGGCTGGCACTCTGGAATGATGATGAGAGAGAGAGAGAG 990
 QY 931 CAGAGACCCACACACTGGGCTGAGGCTTGGAAACAATGATGAAGAGAGAGAG 990
 Db CAGAGACCCACACACTGGGCTGAGGCTTGGAAACAATGATGAAGAGAGAGAGAG 1030
 QY 955 AAGATGAGGAGTGAAGTGAACAAGGAGGCAATGAAGTGTCCCTCTCAGAGAGAG 1014
 Db AAGATGAGGAGTGAAGTGAACAAGGAGGCAATGAAGTGTCCCTCTCAGAGAGAG 1050
 QY 1015 CCCTATCTGTTGAGTACAGAAACAGCAACTGAGCAGAGTGCAGAAACCCAGAG 1074
 Db CCCTATCTGTTGAGTACAGAAACAGCAACTGAGTGCAGAGTGCAGAAACCCAGAG 1110

QY 1075 AAGCTTGAGTAAAGAGAGAGCGAAATCAGATGCTCTGACAAAGAGAGCCCTCCGATACG 1134
 Db CAGCACAAG 1170
 QY 1135 GTTCCACCTCAGTTCATAGGCTGAAGCCGGCTGACATCAACGTAAATGAGAGCCCTG 1194
 Db ACTCCACCTCAGTTCATAGGCTGAAGCCGGCTGACATCAACGTAAATGAGAGCCCTG 1230
 QY 1195 GACTCTCAAGGAG 1254
 Db GACTCTCAAGGAG 1290
 QY 1255 ACTCAGTACAG 1314
 Db ACTCAGTACAG 1350
 QY 1315 ACCCTGGCAGAACCTCTCCGGGAATTCACCCCTTCCCTGAAGAGGCTTCTGTGGCACT 1374
 Db ACCCTGGCAGAACCTCTCCGGGAATTCACCCCTTCCCTGAAGAGGCTTCTGTGGCACT 1410
 QY 1375 GGGAAAGAAACAGTCTTAATGCTTCTTAACAGAGCTGTGGAGAGAGAGAGAGAG 1434
 Db GGGAAAGAAACAGTCTTAATGCTTCTTAACAGAGCTGTGGAGAGAGAGAGAGAGAG 1470
 QY 1435 GATCTACCTGTCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1494
 Db GATCTACCTGTCCAGGAG 1530
 QY 1495 TTTCAAG 1554
 Db TTTCAAG 1590
 QY 1555 TGCATGATGTGCTCAATATTTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1614
 Db TGCATGATGTGCTCAATATTTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1650
 QY 1615 GATATCTCATGAG 1674
 Db GATATCTCATGAG 1710
 QY 1675 GTCAACTGAG 1734
 Db GTCAACTGAG 1770
 QY 1735 TCTGTGTCTCTGTGTCTGAAAGTTGATGATTAATCAACAGAGAGAGAGAGAGAG 1794
 Db TCTGTGTCTCTGTGTCTGAAAGTTGATGATTAATCAACAGAGAGAGAGAGAGAGAG 1830
 QY 1795 GAATTCAG 1854
 Db GAATTCAG 1890
 QY 1831 GATATTAACAAG 1890
 Db GATATTAACAAG 1930
 QY 1855 ACAAGGAG 1914
 Db ACAAGGAG 1950
 QY 1915 ACCTGGAG 1974
 Db ACCTGGAG 2010
 QY 1975 AAGTCTCACTCCGAG 2034
 Db AAGTCTCACTCCGAG 2070
 QY 2035 AAG 2094
 Db AAG 2130
 QY 2095 CCGTTTCTGAG 2154
 Db CCGTTTCTGAG 2190

2155 GACAGAGCCCTTCTGCTTGCACCTTACCTGAGTGATGCTGAGACCTGACAGATC 2214
2191 GAGAGGAGCCCTTCTGCTGCTGCTCCACCTGAGTGATGCTGAGACCCGAGATGTC 2250
2215 CAAGTTGTGCTGCTCTGAGGAGATTTCTGACCGCTGAGCATGGAATTTGGCTCCAAACA 2274
2251 CGTGTGTGCGCGGTGGGGGATTTCTGACTGCTGGCAGTGAATTTGGCTCCAAACA 2310
2275 GAGAGCTCCCGCATGTCACACACAGATTCGGGGGAGTGTCACTACAGTGACGAGGGGAC 2334
2311 GGGAGCTCCGCTGATGTCATCACAGATATCGGGAGCTGTCTCTATGTTACGAGGGGAC 2370
2335 GGGCTCCGAGAGATGTCACACCTTATCTATATCTTCGGAGTTTAAACAGAACTC 2394
2371 GGGCTCCGAGATGTCACACCTTATCTATATCTTCGGAGTTTAAACAGAACTC 2430
2395 ACAGGCTACCGCGTGGGACCGGCTGATGCCAATGACAGAAATGCAATTTCTCATACAGCT 2454
2431 ACTGTTACCGCGGTGGGACCTGATGCCAATGCAATGCAATTTCTCATACAGGCTC 2490
2455 GTTCCCGAGCAAGGCTGAGGATCTTATGAGCGCAATCCAACTCTGATGAGAGATG 2514
2491 GTTCTTGAGAGAGGCTGAGAACTTATGAGCCAGCTCAAGACTCTGATGACAGAGATG 2550
2515 AAGATGATCATGAGTAAATTTTCATGAGAGCTGAGAGGTCATCAAGTGTGATGGA 2574
2551 AAGATGATCTGAGTAAATTTTCATGAGAGCTGAGAGGTCATCAAGTGTGATGGA 2610
2575 GGCAGCATTTTATGTCATCTGACAGATTTGCAATCTGATTTTGTGACGCCAATTTGTT 2634
2611 ACCAATATTTGTGCAACCACTGACAGATTTGACCTGATTTTCAAGCCCACTTTTTC 2670
2635 GACCATCTCGGCAATGCTTGGAGAGTCTGATGAGAGGAGGCCAGAGCTCTGATCAAC 2694
2671 AACCATCTCTCAATGCTTGGAGAGTCTGATGAGAGGAGGCCAGAGCTCTGATCAAC 2730
2695 CTCGTGACTTCTGAAACCCCACTATCATGCGAGGTGTCTCTGGGAAACCAAGACAG 2754
2731 CTGTGTGACTTCAAGAAACCCAGATATCATGCGAGGTGTCTCTGGGAAACCAAGACAG 2790
2755 TGCCAGTGCAGAGGCGCGGCTTTTGTGATGCTGCTTGCACCTGCGGAGAACTCC 2814
2791 TGCCCGTGCAGAGGCGCGGCTTTTGTGATGCTGCTTGCACCTGCGGAGAACTCC 2850
2815 CAAGAGTAGTGCAGAGGCGCGGCTTCAAGCGGAGCTTCAAGAGAGATGCGGAGCTG 2874
2851 TATGATGCGGAGAGAGAGGCGGCTTCAAGCGGAGCTTCAAGAGAGATGCGGAGCTG 2910
2875 GTGGGATCAGGCGCGTATGACAGCGAGAGAGCTTCTGTGTGCTGACGCCCTTCTTC 2934
2911 GTGGAATCAAGGCGCGTATGACAGCGGAGAGAGCTTCTGTGTGCTGACGCCCTTCTTC 2970
2935 CAGAACTCAGGCTTCTGTCTGTGCGGAGTGGGCTCCAGATACGTCCTTTTGGCCCA 2994
2971 CTCAACATCCGCGCTCCCATCTGAGAGATGGGCGTCAGACACTCTCTTTTGGCCCG 3030
2995 GACTGCTACCAACCAATTCAGAAATTCACCTCCAGCGTGGCAGAGCCCTTTGAGCAAT 3054
3031 GACTGCTACCAACCAAGGCGAGAAATTCACCTCCAGCTCTCTAGAGCCCTTTTGGTCAAT 3090
3055 ATGCTTGAACCACTTGAAGCAAAACAGAGACCTTGAACCTGAGAGAGAGATGCCATC 3114
3091 ATGCTTGAACCTTGAAGCAAAACAGAGATGATCTGTGACCTGACAGAGATCTCTCCCG 3150
3115 ACTGTCTCACTCAAGATGAGCCCTTCTGAGAAACCTTCCGAAATGTAATCAACGTAAC 3174
3151 CCTGCCCCACCCAGAGAGAGCCCTTCTGAGAAACCTCCCAAGACGAGATCAACGTAAC 3210
3175 CCCATCAAGCAAGCATGAGAACTGGGGAGTGAATCTTCTGTGATGAGAGATGAGGCT 3234
3211 CCCACCAAGCAAGCATGAGAACTGGGGAGTGAATCTTCTGTGATGAGAGATGAGGCT 3270
3235 TCCATATGATTTCCAACTCTGTCCACAGCTCCGACAGAGATCAATCAAGTGGTGCC 3294

3271 TCCATATGCTCTCCCACTCTGATTTCAACAGCTCAACAGCAACATCAAGTGGTGCA 3330
3295 GCCCTGGGATGCTCTGATCTACAGAGAGTGGAGCTGACCAACAACTCCAGTACCTA 3354
3331 GCCCTGGGATGCTCTGATCTACAGAGAGTGGAGCTGACCAACAACTCCAGTACCTG 3390
3355 CCCACATCTTGAAGGAGCTCTCTTGAAGCATTTGAGGGATGGAATTTGAGATCTAC 3414
3391 CTCATGCTCTGAGAGGGGCTCTCTGAGCATTTGAGAGATGAGAGCTTACAGACCCAC 3450
3415 ACCAATGCTCCCAATTTGAGAAAGTTCACCTTACTCTTGGCTTCTTACACAC 3474
3451 ACCAATGCTCCCAATTTGAGAAAGTTCACCTTACTCTTGGATTTCTTACATGCT 3510
3475 ACCCTGGAGGGGACAGAGAGCTAAATGTCAGAGCGGAGAGGGGACAGAGTACGAGCATG 3534
3511 ACCCTGAGAGAGAGGAGATTTACGTCAGTGGAGAGAGAGAGAGAGCTCGGAGCATG 3570
3535 CAGAGCCAGGCTGAGGACCTGATGAGAGAGTAAACAGCCCGACATCAACCTGAG 3594
3571 CAGAGCCAGGCTCGAGAGCTGATGAGAGATGAGAGAGAGAGAGAGAGCTCGGAG 3630
3595 AAGATGAGAGTGTGTCATCTTCTTATTTGGGTCAAGCATTTGTGTCTTACTGTAG 3654
3631 ATGAGCTGAGAACTGATCACTCTTCTTATTTGGAGAGCATATCTGTGTCTTACTGTAG 3690
3655 AATCCGAGGCGCACTTGGGACAGGAAATATGTTACAGACATCCAAAGGCGCTGAGCATC 3714
3691 AATCCGAGGAGCACTGCGGAGAGAGATATGTCAGACATCCGAGAGGCGCTGAGCATC 3750
3715 CTCCTGAGAGAGCTCCCAAGGCTTTTGTGTCAGAGTGTGAG--GCTATGAGAGTGGCT 3771
3751 CTGATGAGAGAGCTTCCGAGGCGCTTCTCATCAAGTATGAGATCATGAGAGTGGCT 3810
3772 AGCTGTACAGAGGCGCAAGGCGGAAATG--TGCATGCTGCGAGCTCAGAACATCTGC 3828
3811 GGCCTGACAGAGGCGCAAGGCGGAGATGTCAGCTGCTGTCAGAGACAGAGACCTGC 3870
3829 ACTTGCCTCAGACATCGGCAAGCTCCCTGAGAGAGAGAACTGAGAAAGTGAATCG 3888
3871 AGTGTCTCCGACATCTTCCAGCTCTCTGTGTATCCAGAGCTGAGAAAGTGAATCG 3920
3889 AACCTCAGAGATGAGATCTCAGATTTCTCTTACTGAGCAACATATACAGAGCTGAGAG 3948
3931 AACCTCAGAGATGAGATCTCAGATTTCTCTTACTGAGCAACATATACAGAGCTGAGAG 3990
3949 TTTGCGGTGTGTGAGAGCTTTTCTTCAAAACACACTCACCCCACTGAGAGAGAGG 4008
3991 TTTGCGGTGTGTGAGAGCTTTTCTTCAAAACAGCTCAATCCGCTGAGCAACAGCTGCG 4050
4009 GACATGACCTCACTCTTCTTCTCCGAGAGCTTTTCACTTCTGACACCGGCGGATGCC 4068
4051 AGCACTACCCCACTCTTCTTCTGAGAGAGCTTTTCACTTCTGAGAGAGCTGAGAG 4110
4069 GAGATGCGCATGAGATCTGAGAAACATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 4128
4111 GAGATGCGCATGAGATCTGAGAAACATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 4170
4129 AACAACTTCAACCAAGCGGAGCAAACTCAAGTGGCTTCTCTGAGAGAGCTTACCTC 4188
4171 AACAACTTCAACCAAGCGGAGCAAACTCAAGTGGCTTCTCTGAGAGAGCTTACCTC 4230
4189 TACACCTGCGGAGAACAGCCCATTTGCTCCAGACCAAGGCTGAGAAAGGCGCCGAGAGTGC 4248
4231 TACACCTTCAAGAACACTTGTCTCTCCAGTCCAGCTGAGAAAGGCTCTGAGTGTAGCC 4290
4249 TACTGAGCTGTCCAGTGTGAGAGCGGAGTGGCTTGTGTGTGAGATCATG 4300
4291 CCCGAGATGTGTAGAGAGAGAGAGAGAGTGGCTCTCTGAGTGGCTCATAG 4342

RESULT 5

D63648 LOCUS D63648 4590 bp mRNA linear ROD 01-NOV-2000
 DEFINITION Rattus norvegicus mRNA for phospholipase B, complete cds.
 ACCESSION D63648
 VERSION D63648.1 GI:2696235
 KEYWORDS phospholipase B.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 REFERENCE 1 (sites)
 AUTHORS Takemori, H., Zolotarov, F. N., Ting, L., Urban, T., Komatsu, T., Hatano, O., Okamoto, M. and Tojo, H.
 TITLE Identification of functional domains of rat intestinal phospholipase B/lipase. Its cDNA cloning, expression, and tissue distribution
 JOURNAL J. Biol. Chem. 273 (4), 2222-2231 (1998)
 MEDLINE 98113187
 PUBMED 9442065
 REFERENCE 2 (bases 1 to 4590)
 AUTHORS Tojo, H.
 TITLE Direct Submission
 JOURNAL Submitted (22-JUL-1995) Hiromasa Tojo, Osaka University Medical School, Molecular Physiological Chemistry, 2-2 Yamada-oka, Suita, Osaka 565, Japan (E-mail: htojo@mr-bio.med.osaka-u.ac.jp, Tel: 06-879-3283 (ex. 3283), Fax: 06-879-3288)
 FEATURES
 source location/Qualifiers
 1..4590
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="SD"
 /db_xref="taxon:10116"
 /tissue_type="small intestine"
 36..4388
 /codon_start=1
 /product="phospholipase B"
 /protein_id="BA23813.1"
 /db_xref="GI:2696236"
 /translation="MESWPGVSLVGLLLLLLLGQSPQIHSSGENTSPQVFRFLK
 NESFCKPKLESLVLSKSVSLRPDIKLVAILGNLETPPAGSGVMNEKQSL
 ELONVIGIMTALSDIIRHNPSPVLMPTCSGPGTGAHTIADLIIOAKELVRHLD
 NPELDKDKMLITVLFNSOCHLCSDDCKRHLKMKHMSGLVDYLRERAEV
 NLVDSEVLTMAOHOETGSPAPBEICSEETIKLSKAMQMSYOEAMDELAASKE
 NGETRAVYVQSFSEVEPLERPSPODSTTLARITNSMSEVGRKDTLNEKRT
 MKGQSEPLFTYRNSNTQARQLKPTGRQMEGRFTCPDPSDLSPTVHRLRP
 ADIKVIGAMDLSLTAAGSSPGNVLDVLTQYRGLSWSVGBETLETTLANILRE
 IFVFIGNDLGSNNLAFSPOTFDNITKTLADILHAEPVAFVNMVSEITPLR
 ELNEBKVGCPMILRSICPCVNLGENSELAOLVERNNOYSEATKLESGRYDTR
 DDETVVLQPMFENVNPRILEGLPDSFPAPDCTHRNVKTHASALAKMKMLPYGR
 KTHQNEIKPIWCPNQTSPLSTTKNSLGHGTSMSCERKPSAPSPSVTLTPA
 DIQVVALGDSVTAAGNISQEDGLADVTQYRGLSVSAGDEFLANVTTLPIILRE
 NGVLTAVGTGVNVSASAFVLAQVGAKEENLASOVQTLIQKNDITRVNPDQK
 ITVMIGASDLCECDNSRYSAANFSDHLENALDIHKVEPRALVNLVDMPNSIIR
 VLEKNDKCPVNOTSVLNCVLTFRGDSHLEALAEATKESYOSMOLVESGYDRE
 DRSVYQPLFNRILRPLENGANDTSPFADICLLSOKHTQALRLMANMELPKK
 MDLDELKELALACPTKDKPLRTFRNSNTYPIKPAIEWGSDFLCTEGSPSKVPT
 SVHELPSDLKVVAAAGDFLTATGARPSSSSILDTPIWRGLSISIGDDELEHTLP
 NLIKFNPSILIGSTGLTENTAGLAEBAADMPAOQALVKMKSPPTINIOED
 NKITLLIGNNDCLYCEPDENYSTREYVYIOHALDIPEELRPVFINVEYMEISG
 LHHDOGKAMPVLAOKNGSCIKRSIONAMOBELKXVNGALSSELSYHRYMRE
 DRAVYVQPPFRNTFVPLDERGGDLTFESDDCFHFSYRGAENAILMANMELPVGK
 TTSNFTYNTKTKCPSPENPLTYVRNSQILLDKAKENSNTLYMAVPAVAGGLVVG
 IILQMMLRTVRLVQ"

BASE COUNT 1195 a 1243 c 1142 g 1010 t
 ORIGIN

Query Match 61.4%; Score 2687; DB 10; Length 4590;
 Best Local Similarity 77.1%; Pred. No. 0;
 Matches 3340; Conservative 0; Mismatches 970; Indels 21; Gaps 5;
 QY 28 CTGAGCTGCTGCTTCTGGGGCAAGGACCCCTCATGATCCATCCTCTCTAGAAAG 87

Db 69 CTGCTGCTGCTACTAGCTGGGCAAGGCGCCCTCCCAATCCACGGCTCTTCGAGAA 128
 QY 88 AGTACATTGAAGGAGAGCTATGCGCAGAGACCCCTAAGAAATTCATTCCTCATGAC 147
 Db 129 AACACATCGAGGCCCGCAGCAAGTGTTCGACCCCTGAAGAAATTTTCATTCCTTCGAC 188
 QY 148 CCAATAAATTAGAGTGAATATGCTCTTAATTCAGTTCACTCTCGAAGCCTTCGAT 207
 Db 189 CCAGAGATGTAACATGACTGTGCTTTCTAATGTCAGTTCACTCTCGAAGCCTTCGAC 248
 QY 208 ATTAAATTGTGCGACCCATTTGCAATCTGGAATTCCTCGAGACCCGAGGAGCGCG-- 265
 Db 249 ATTAACTGTGCGACGACATCGGCAACTGAAATCTTCAGCCCTCGCTCAGGCGTG 308
 QY 266 ---ATCTGAGAGCAAGCAAGTGAAG--GCCACAGAGGTGTCATGAGAGTG 318
 Db 309 GTCAATGAGAGAACCTCAAGCCTCGAGAGGAACTACAGAAATGTGTCATGAGAAATC 368
 QY 319 ATGACAGCTCTTCAGACATCATCAGATTAATTCAGTCTCTGTCGAATGCTGTGTC 378
 Db 369 ATGACAGCTCTTCAGATCATCATGACATTTCAACCTCTCTGTCGATGCCCAGCTGT 428
 QY 379 CACACTGAAAGAGAGTCAATACCCCA--GATGTCTGGAAGACTTGTGATTCAGCT 435
 Db 429 TCTCTGGAGAGGTACTGAGGCGCACACTATTGCGGAAGACTGTGATTCAGGCT 488
 QY 436 CAAAGACTGTGAGAAACATGAAGAGAACCTTGCACTTGACATTTGACTGGAAG 495
 Db 489 AAGAGCTGTGAGGATCTGAAAGCAACCCGGAATTCGATTTTGAGAGAGCTGGA 548
 QY 496 CTGATGATGTGTTCTTCAATGAGAGGAGTGAAGCTGAGCTGAGCTGCTGCTCAACAG 555
 Db 549 CTGATGCTGTGCTTCTGATACAGAGGAGGAGGCACTGTGTTCTTGTATCAGAG 608
 QY 556 AATGCTGTCGCGCGCGCGCGCTGATGAGCTGATGGGGTGTGCACTGACAGAG 615
 Db 609 AAGAGCACTGATGAGAGCAATGAGAGAGTGTCTGAGGGTGTGATTCAGCACTGT 668
 QY 616 GAGTCCCAAGAGCACTTTGTAACCTGTGAGACTCTCTGAGGTGACAGGTTCTGCT 675
 Db 669 GAGTCCCAAGAGCGTTTGTGAATTTGTGAGATCTCTGAGGTTTGAACATGAGCTG 728
 QY 676 GATATCAAGGAGCTTGTGCTGAGCCCTGACAGAGACCCGTGAATTTCTGAGAGAG 735
 Db 729 CAGCATCAAGAGCTGTTCACCTGACAGCAAGAAATTTGCAAAATGCTCAGAGAAATA 788
 QY 736 ACCGCTGCGCCAAAGTGTGATGAGCTGATCTTATCAGAGACCTGGAACAGCTCTG 795
 Db 789 ACGAATTATCAAGGCTGTATGAGAGTGTCTTATCAGAGACCTGGAAGATCTCTG 848
 QY 796 GCTTCAGAGAGTACAGTACAGAGAGTCTTCAAGCTGTTTTCAGCTTTCTTAT 855
 Db 849 GCCTCAGCAGAGTCAATTAAGCATGAGACCTTGCCTGTGTTTTCAGTCTTCTCT 908
 QY 856 GAGACCAACCCATCTTACCTGAGAGACCCCGCACTCAGAAATTTACAGCTGCGC 915
 Db 909 GAG-----GTAAACATACCTTGTGAGAGGCTTCCGCTCAGGATTCACAGCTGCGC 962
 QY 916 TGCACTCTGGAATGATGATGAGAGCCGACAGAGAGAAATGAGCACTGAGTGA 975
 Db 963 CTGAGATCTGGAATGATGATGAGAGCAAGTGTGAGAGATGAGACACTCAATGAA 1022
 QY 976 AAACAGGAGGCAATGAAGTGTCTCTCAGAGAGCCCTATCTGTTCACTACAG 1035
 Db 1023 GCAGAGAGAAACATGAATGTCTCTCAGAGAGCCCTATCTGTTCACTACAG 1082
 QY 1036 AAGAGCACTACCTGACAGCTGACAGAAACCCCAAGACAACTGTGAGAGAGAG 1095
 Db 1083 AATAGCACTACAGGCGCAGAGAGCTGAAACCATGAGAAAGTTTGAAGTGAAGAG 1142
 QY 1096 GCGGAATCAGATGTCTTCAAAAGAGCCCTCGATACAGGTTCCACTCAGTTCAAG 1155

QY	3316	ACAGCAGTGGAGCTCGACCAAAACAATCCAGTGCATCAACCAATCTTGGAGGGGAACTC	3375
Db	3363	ACAAGCCACCGGAGCTCCACCAAGTGAAGTCCAGCACTCTAGACACCCCTGGAGGGGGCTG	3422
QY	3376	TCTTGGACATTTGGAGGGGATGGGAATTGGAGACTCACACACACTGCGCAATCTG	3435
Db	3423	TCTTGGAGCATTTGGAGGAGATGGAAAGCTTGGAGACCCTATACCACTCTCCAAATCTCTG	3482
QY	3436	AAGAAGTTCAACCTTACCTCTGCTTGTCTACAGCACTCTGGAGGGGACAGCAGGA	3495
Db	3483	AAGAAGTTCAACCTTCCATCTCTGGATTTCTCACCCGGTACCTTGGAAACACGCGCAGGA	3542
QY	3496	CTAATATGTGACGCGGAAAGGGGCGCAGAGCTTAGGGACATGCCAGCCAGAGGCTTGGGAACTTG	3555
Db	3543	TTAATATGTGGCAGAAAGAAAGGGCGCAGAGCTCAGAGACATGCCGCGCAGAGCTCAAGCCCTTG	3602
QY	3556	GTAAGCGCAATGAAAAACAGCCCGGACATCAACTGGAGAAAGACTGGAAAGCTGGTCA	3615
Db	3603	GTGAAGAAGATGAAAAAGCACCCTTACATCAACATACAGAAAGACTGGAACCTGATTAACA	3662
QY	3616	CTCTTCAATGGGGTCAACGACTTGTGTCATTAAGTGAATCCGGAGAGCCCATTTGGCC	3675
Db	3663	CTCTCAATCGGGAACAAAGACCTGTGTCTTACCTGTGAGATCCGGAAATCTACTCAAC	3722
QY	3676	ACGGAATATGTTCACGACATCCACAGGCGCTGGACATCTCTCGAGAGACTCCCAAG	3735
Db	3723	AGGAGTATGTCAAGTACATCCAGATCCCTTGGACATCTTCTATGAGAGACTTCCAGG	3782
QY	3736	GCTTTCGTCAACGTGTGTGAAGTCAATGAGCTGTAGCTGACCTGTACCGAGGCCAAGCGGG	3795
Db	3783	GTTTTCATCAACGTGTGTGAAGTCAATGAGCTGTCCGGTGTCTTCAACGACCAAGGCGGG	3842
QY	3796	AAATGTGCGCATGC---TGGCAGCTCAGAACAACTGCAACTTGGCTTCAGCACTCCCAAGC	3852
Db	3843	AAATGTGCGCATGCCTGTGGCTGTCCAGAAAACTGCAATTCCTTTAAACGCTTCAAAAC	3902
QY	3853	TCCCTGAGGAAGCAAGAACTGAAGAAATGAACTGGAACCTTCCAGATGAGCATCTCCAGT	3912
Db	3903	CTCATGTGGAATGCAAGAGACTGAAGAAATCAACGGGAACCTTCAAGTGGCCCTCTCGGAG	3962
QY	3913	TTCTCTCTACTGGACAATATACACAGCGTGAAGACTTTTGGGTTGTGTGACAGCTTTC	3972
Db	3963	CTCTCTCTACTGGACATCGGTATCATGACGCTGAAGGACTTTGCAAGTCACTGTCAACCTTTC	4022
QY	3973	TTCCAAAAACAACATCAACCCCACTGAAGAGAGGGGGAACATGAACCTTCACTTCTTCTCC	4032
Db	4023	TTCCGGAATACCTTTGTCTCCATGGATAGGCTGGGGGCGCTCGAATCTACCTTCTTCTCT	4082
QY	4033	GAGGACTGTTTTCATCTTCTCAGACCGCGGAGTGGCCGAGATGGGCATCTGCACTCTGAAAC	4092
Db	4083	GAAAGCTGTTTCCACTTCTCAGTCCGTGTGGGATGTGAAGTGGCATTTGGCCCTCTGGAAC	4142
QY	4093	AACTATGCTGGAACCAATGAGCGCGCAAGCTACCTTCAACAACTTCAACCAAGCGGAGCC	4152
Db	4143	AACTATGCTGGAACCAATGAGCGCGCAAGCAACATCTCAATTAATCTTCAACCAACCGGAGCC	4202
QY	4153	AAACTCAAGTCCCTCTCTCTGAGAGCCCTTAACTCTTACACCTCTGGGAAACAGCGGATTG	4212
Db	4203	AAACTCAAGTCTCCTCGCTGAAAAACCTTTCTCTTACACTGTCTGGAAACAGCTCAATT	4262
QY	4213	CTCCAGACCAAGCTGAAGAGCCCGCAGAGTGTCTACTGGGCTGTCTCCAGTGGACGG	4272
Db	4263	CTTCTAAGCAAGGCTTAAGAAAACTCCATACCTTACTGTGGGCACTGCCAGTGGCTGCA	4322
QY	4273	GGAAGTCCGCTTTGTGTGGGCAATCATCCGGAACAGTGTCTGAGGTGCAGAGAGAGTGGC	4332
Db	4323	GTAGGTGCCCTGTGATGTGGCATCTTGGAAATGATGTGTGGAGAACCTGTGAGACTCGTC	4382
QY	4333	CGAGAGGAGAGA 4343	
Db	4383	CAATATGAAGGA 4393	

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
EL1935	Rat mRNA for phospholipase.	EL1935	EL1935.1	GI:3252702	JP 1997248190-A/1.	
REFERENCE	1 (bases 1 to 4613)	Tojo,H. and Hasegawa,A.				
AUTHORS	NEW PHOSPHOLIPASE AND DNA CODING THE SAME					
TITLE	Patent: JP 1997248190-A 1 22-SEP-1997;					
JOURNAL	TUO HIROMASA, TONEN CORP					
COMMENT	OS Rat1us sp. (rat)					
	PN JP 1997248190-A/1					
	PD 22-SEP-1997					
	PF 15-MAR-1996 JP 1996086022					
	PI TUO HIROMASA, HASEGAWA AKIRA					
	PC C12N15/09, C07H21/04, C07K14/47, C12N5/10, C12N9/16, PC					
	C12N9/20//A61K38/46,					
	CC (C12N5/10, C12R1:91), (C12N9/16, C12R1:91), (C12N9/20, C12R1:91);					
	CC strandedness: Double;					
	CC topology: linear;					
	PH Key					
	Location/Qualifiers					
FEATURES						
source	1..4613					
	/organism="Rattus sp."					
	/mol_type="genomic DNA"					
	/db_xref="taxon:10118"					
	/product="phospholipase".					
BASE COUNT	1218 a	1243 c	1142 g	1010 t		
ORIGIN						
Query Match	61.4%	Score 2687;	DB 6;	Length 4613;		
Best Local Similarity	77.1%	Pred. No. 0;				
Matches 3340;	Conservative	0;	Mismatches 970;	Indels 21;	Gaps 5;	
28	CTGAGAGCTGCTGCTCTTCTGAGGCAAGGGGACCCCTCGATCCATACCTCTCTAGAAAG	87				
69	CTGCTGCTGCTGCTGCTGCTGAGGCAAGGGGACCCCTCTCCAAATTCACGGCTCTTCTGGA	128				
88	AGTACATTTGGAAGGCAAGCTATGAGCCAGAGACCCCTGAAGAAATTCATTCACATGCAAC	147				
129	AACACATGAGAGCCCAAGCAAGGTGTTTCGACCCCTGAAGAAATTTTCAATTCCTTGCAAG	188				
148	CCAAATAATTTAGAGTGAATATATGCTTCTTAATCAAGTTCATCTCTGTAAGCCTTGAT	207				
189	CCAAAGAGATTGAGAACTGAGTGTCTTCTTAAGTCAGTTCACTCTCTGAGACCCCTCAAC	248				
208	ATTAAATTTTGGAGAGCATTTGGCAATCTGAAATTCCTCCAGACCCAGGAGCGGGCGG--	265				
249	ATTAAATCTGTTGAGAGGCACTGAGAACTCTTCCAGGCCCTCTGAGCTGAGGCTG	308				
266	-----ATCTGAGAGAGCAAGCTGAGCTGAAG--GCCACAGCAGGTGTGATGGAGTG	318				
309	GTCAACATGAGAGAACTCAAGAGCTCGAGAGCGCACTACAGAAATGTGTGATAGAAATC	368				
319	ATGACAGCTCTTTGAGACATCATATAGATATTTTCACTCTTCTGTTCCAAATGCTGTGTC	378				
369	ATGACAGCCCTTTGAGATATCATGACATTTCAACCCCTTCTGTTCTGATGCCCCACGCTG	428				

[illegible]

Db	1503	AAGCTGTGAGCCTGTGATGAAGATGCAAGACATATACTTTGAGGAAGACTGGAAATATA	1562
Qy	1516	ATAACCCCTGTTATATAGAGCGGCATATGACCTCTGTGATTTCTGCAATGATCTGTCTCACTAT	1575
Db	1563	ATCATCTGTGTTTATATAGAGGCAACAGACCTCTGTGGCTCTCGCATATACTGGCTCGCTTT	1622
Qy	1576	TCCTCCCAAGAACTTCCACAGACAAATCATGAGCGCCCTGTGACATCTCTCCATGCTGAGCTT	1635
Db	1623	TCCTCCCAAGAACTTCCACAGACAAATCATGAGCGCCCTGTGACATCTCTCCATGAGAGT	1682
Qy	1636	CCTGGGCAATTTGTGACCTGTGAGCGGTGCTTGATGTCATCCCTGAGGAGCTGTAC	1695
Db	1683	CCCCGGGCTTTGTGAAACATGGCTCTGTGGATTTGATATACCCCTTGAGAGAACTGTTC	1742
Qy	1696	CAGAGAAAAAAGTCTACTGCCCCAAGATGATCTCAGGCTCTGTGTCCCTGTGCTTG	1755
Db	1743	AATGAACCTTAAGTCACTGCTGCCAGGATGATCTCTGAGAGCCTGTGTCTTGTGTCTTG	1802
Qy	1756	AAGTTGATGATTAACATCAACAGAACTTGCTACCTCATGCAATTCACAAAGATTTTCAG	1815
Db	1803	AACCTTGGTGAAGAACTCAGACAGAACTTGCCCAACTGTGGAAGAAACAGGCAAGTATCAG	1862
Qy	1816	GAGAAAGCCACCACTGATTTGAGGTGGGCGCATATGACCAAGGGAAGTTTATCTGTG	1875
Db	1863	GAGAAACTGGAAGAACTGATTTGAGAGTGGCCGATACGACACAGGAGTGAATTCACCGTG	1922
Qy	1876	GTTGTGACAGCCGTTCTTTGAAAAACGTGACATGSCAAAGACTCTCGGAAGATTTGCTGAC	1935
Db	1923	GTCTCTCAGCCCATGTTTAAATGTCTGATGTCATGCCAGAGCCCTGGAAGGCTTGGCCGAC	1982
Qy	1936	AACCTTTTCTTGCTCTGCTGACTGTTTCACTTCAGCAGCAAGTCTCATCTCCGACAGCC	1995
Db	1983	AACCTTTTCTTGCTCTGCTGACTGTTTCACTTCATGTCAGAGACTCAGCTCTGCTCAGCC	2042
Qy	1996	AGTGTCTTTGGAACAATATGCTGGAGCCTGTGTGGCCAGAAAGCACTGCTCATTAAGTT	2055
Db	2043	ATGCGCCTCTGGAAGAACATGCTAGAACCTGTGGGCGGCAAGACATCAGAAATTTT	2102
Qy	2056	GAAGAACAGATCATATATCATATGTCGGAACCAAGTCAAGCCGTTTCTGAGACCTTACAG	2115
Db	2103	GAAATCAGAGTCCCTATATCATGTGTCCAAACCAAGCCTCACGTTTCTGAGACCAACAG	2162
Qy	2116	AACAGATGCAAGGTCTATGGAAGCTTGCTGCCATGACAGGACAGAGCCCTTGTGCTTG	2175
Db	2163	AACAGAACTTGGGACATGGAACCTTGATGTCTTGTGAGGAGAAAGCCCTCTGCTCA	2222
Qy	2176	CACCTTACTCATGTGATATGCCCTGAGACCTGCAAGACATCCAAATGTGTGCTCTGAGG	2235
Db	2223	CCACCAACTCATGTGATATGCCCTGAGACCTGCAAGACATCCAAATGTGTGAGCTTGGGA	2282
Qy	2236	GATTCCTGACCGCTGGCAATGGAATTTGGCTCCAAACAGAGCACTCCCGATGTCAAC	2295
Db	2283	GACTCTGTACTGTGGCAATGGAATCCAGTCTCCAGAAAGTATCTCGCTGATTTTACC	2342
Qy	2296	ACACAGTATCGGGGACTGTATATCAAGTGCAGAGGGGAGCGCTCCCTGAGAAATGTGACC	2355
Db	2343	ACACAGTATCGAGGACTGTCTTACAGTGTGAGTGGGGAACAAGTTCTCGAGAAATGTGACC	2402
Qy	2356	ACCTTAACCTAATATCTCTGGGAGTTTAAACAGAAACCTTCAAGGCTTACCGCGTGGCAGC	2415
Db	2403	ACCTTGGCCCAATCTCTCCGGGAATTTAAATGGAATCTTCAAGGCTTACCTCAGTCGGAACC	2462
Qy	2416	GGTGTAGGCAAGACAGCAATGATTCATCCATCAAGCTGTTTCCGGAGCAAGGCTGAG	2475
Db	2463	GGTGTAGTCACTCTGCAAGCGGTTCTTTAACAGGCTGTCCCTGGGCAAGAGCTGAG	2522
Qy	2476	GATCTTATAGGCAAGTCCAAACTCTGATGACAGAAATGAAGATGATCATTAAGTAAAT	2535
Db	2523	AACCTTGAAGTCAAGTCCAGACTCTGATTCAGAAATGAAGATGACACAGAGTGAAC	2582
Qy	2536	TTCCATGAAGACTGGAAGGTCAACACAGTCTGATCGAGGCAAGCAATTTATGTACTAC	2595

Query Match 31.1%; Score 1362; DB 6; Length 1835;
 Best Local Similarity 89.3%; Pred. No. 0;
 Matches 1582; Conservative 0; Mismatches 0; Indels 190; Gaps 3;

QY 2618 CTGACGCACTTTGTTGACATCTCCGCAATGCTTGAAGTCTTCATAGAG----- 2671
 DB 1 CTGACGCACTTTGTTGACATCTCCGCAATGCTTGAAGTCTTCATAGAGCTTT 60
 QY 2672 -----AGTGGCCAGAGTCTGCTCAACCTCGTGAAGCTTCTGAAACCCCACTATGCTG 2725
 DB 61 TCCCTTAGGAGCCAGAGTCTGCTCAACCTCGTGAAGCTTCTGAAACCCCACTATGCTG 120
 QY 2726 GGGAGGTGTTCTGAGAAACCCAGCAAGTGCCTGAGCAGAGCCAGGCTTTGTGTA 2785
 DB 121 GGGAGGTGTTCTGAGAAACCCAGCAAGTGCCTGAGCAGAGCCAGGCTTTGTGTA 168
 QY 2786 ACTGCGTTGACCTTGGGAGAACTCCAGAGCTAGCCAGGCTTGAAGCTTCAAGCC 2845
 DB 169 ----- 168
 QY 2846 GAGCGTACCGGAGAGATCCGCAAGTGGTGAAGCCGCTATGACACCGAGAGG 2905
 DB 169 -----GAGCAGATCGCGAGCTGGGGGTGAGGCGCTATGACACCGAGAGG 218
 QY 2906 ACTTCTGTGAGTCTGAGAGCCCTTCTTCCAGAACTCCAGCTCCCTGTCTGGCGGATG 2965
 DB 219 ACTTCTGTGAGTCTGAGAGCCCTTCTTCCAGAACTCCAGCTCCCTGTCTGGC----- 273
 QY 2966 GAGTCCAGATAGCTCTTCTTGGCCAGACTGATCCACCAATCAGAAATTCAGCT 3025
 DB 274 ----- 273
 QY 3026 CCCAGCTGGCAGAGCCCTTGGACCAATATGCTTGAACCACTTGAAGCAAAACAGAG 3085
 DB 274 -----GCTTGAACCACTTGAAGCAAAACAGAG 302
 QY 3086 CCCTGAGCTGAGAGAGAGATGCCATCACTGTCCCACTGAGAAATGAGCCCTTCTGTA 3145
 DB 303 CCCTGAGCTGAGAGAGAGATGCCATCACTGTCCCACTGAGAAATGAGCCCTTCTGTA 362
 QY 3146 GAAACCCCTCGGAATAGTAACTACAGTACCCCATCAAGCCAGCTTGAAGATCGGGCA 3205
 DB 363 GAAACCCCTCGGAATAGTAACTACAGTACCCCATCAAGCCAGCTTGAAGATCGGGCA 422
 QY 3206 GTGACTTCTGTGTGACAGAGTGAAGGCTTCCATATGTTCCAACTCTGTCCACAGC 3265
 DB 423 GTGACTTCTGTGTGACAGAGTGAAGGCTTCCATATGTTCCAACTCTGTCCACAGC 482
 QY 3266 TCCGACAGAGAGATCAAAAGTGTGCGCCCTGGGTGACTCTCTGACTACAGCAGTGG 3325
 DB 483 TCCGACAGAGAGATCAAAAGTGTGCGCCCTGGGTGACTCTCTGACTACAGCAGTGG 542
 QY 3326 GAGCTGAGCAAACTCCAGTGAAGTCCCACTATGTTGAAGGGAAGCTCTTGGAGCA 3385
 DB 543 GAGCTGAGCAAACTCCAGTGAAGTCCCACTATGTTGAAGGGAAGCTCTTGGAGCA 602
 QY 3386 TTGAGGGGATGGGAATTTGAGACTCAACCACTGCCCCACATTTCTGAAGAGTTCA 3445
 DB 603 TTGAGGGGATGGGAATTTGAGACTCAACCACTGCCCCACATTTCTGAAGAGTTCA 662
 QY 3446 ACCCTTACCTCTTGGCTTCTTACAGCAGACTGGGAGGGAAGCAGAGCACTAAATGCG 3505
 DB 663 ACCCTTACCTCTTGGCTTCTTACAGCAGACTGGGAGGGAAGCAGAGCACTAAATGCG 722
 QY 3506 CAGCGAAGGGGAGAGTGAAGTCACTGAGCCAGGCTGAGCACTGTGAGAGGAA 3565
 DB 723 CAGCGAAGGGGAGAGTGAAGTCACTGAGCCAGGCTGAGCACTGTGAGAGGAA 782
 QY 3566 TGAAGAAAGAGCCGAGATCACTGAGAGAAAGCTGGAAGCTGTCACTCTTCAATG 3625
 DB 783 TGAAGAAAGAGCCGAGATCACTGAGAGAAAGCTGGAAGCTGTCACTCTTCAATG 842

QY 3626 GGGTCAACGACTTGTGTCACTTGTGAGATCCGAGGCCCACTTGGCCAGGAATATG 3685
 DB 843 GGGTCAACGACTTGTGTCACTTGTGAGATCCGAGGCCCACTTGGCCAGGAATATG 902
 QY 3686 TTGAGCACTCCAAAGGCTTGAAGATCTCTGTGAGAGTCTCCAAAGGCTTTGTCA 3745
 DB 903 TTGAGCACTCCAAAGGCTTGAAGATCTCTGTGAGAGTCTCCAAAGGCTTTGTCA 962
 QY 3746 AGTGGTGAAGGATGAGAGCTGAGTACCTGATACCAAGGAGGCAAGGAGGAAATGTC 3805
 DB 963 AGTGGTGAAGGATGAGAGCTGAGTACCTGATACCAAGGAGGCAAGGAGGAAATGTC 1022
 QY 3806 TGCTGAGCTCAGAACTGCACTGCTGAGCACTGCAAGCTCCCTGAGAGAGC 3865
 DB 1023 TGCTGAGCTCAGAACTGCACTGCTGAGCACTGCAAGCTCCCTGAGAGAGC 1082
 QY 3866 AAGAACTGAAGAAAGTGAAGTGAAGTCAAGTCTCAGATGAGCATTTCTTCTTCTG 3925
 DB 1083 AAGAACTGAAGAAAGTGAAGTGAAGTCAAGTCTCAGATGAGCATTTCTTCTTCTG 1142
 QY 3926 ACCAATACACAGAGGAGGAGTGGAGCTTGGGAGGAGCTTCTTCCAAACACAC 3985
 DB 1143 ACCAATACACAGAGGAGGAGTGGAGCTTGGGAGGAGCTTCTTCCAAACACAC 1202
 QY 3986 TCACCCCACTGAAGAGAGAGGAGAGTCACTCACTTCTTCTCCAGAGAGCTGTTTC 4045
 DB 1203 TCACCCCACTGAAGAGAGAGGAGAGTCACTCACTTCTTCTCCAGAGAGCTGTTTC 1262
 QY 4046 ACTTCTGAGAGGAGGAGTGGAGATGAGTCCATGCACTTGAACCAATCTGTGAG 4105
 DB 1263 ACTTCTGAGAGGAGGAGTGGAGATGAGTCCATGCACTTGAACCAATCTGTGAG 1322
 QY 4106 CAGTGGGCGGAGAGTACTCTCCAACTTCAACCTCCAGAGGAGGCAACTCAAGTGGC 4165
 DB 1323 CAGTGGGCGGAGAGTACTCTCCAACTTCAACCTCCAGAGGAGGCAACTCAAGTGGC 1382
 QY 4166 CTTCTCTGAGAGCCCTTCACTTCAACCTTCCAGAGGAGGCAACTTCCAGAGG 4225
 DB 1383 CTTCTCTGAGAGCCCTTCACTTCAACCTTCCAGAGGAGGCAACTTCCAGAGG 1442
 QY 4226 CTGAAGAGGCGGAGAGTCTTCACTGAGTCTCCAGTGGAGGAGGAGTGGCTTGG 4285
 DB 1443 CTGAAGAGGCGGAGAGTCTTCACTGAGTCTCCAGTGGAGGAGGAGTGGCTTGG 1502
 QY 4286 TGGTGGGAGTCAAGGAGAGTGTGAGAGGAGGAGAGGAGTGGCGGAGGAGGAGATC 4345
 DB 1503 TGGTGGGAGTCAAGGAGAGTGTGAGAGGAGGAGAGGAGTGGCGGAGGAGGAGATC 1562
 QY 4346 CTCGAATGAGCTGCGCACTGTGAGGCTCTAG 4377
 DB 1563 CTCGAATGAGCTGCGCACTGTGAGGCTCTAG 1594

RESULT 8
 BC042674 1790 bp mRNA linear PRI 09-JAN-2003
 LOCUS Homo sapiens, similar to phospholipase B, clone MGC:35447
 DEFINITION IMAGE:5191712, mRNA, complete cds.
 ACCESSION BC042674
 VERSION BC042674.1 GI:27503748
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (02-JAN-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT

Contact: MGC help desk
 Email: cgsaps-x@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nambavali,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC plate: 50 Row: k Column: 10.
 Location/Qualifiers

FEATURES

source

1. 1790
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:35447 IMAGE:5191712"
 /tissue_type="Brain, adult, 6 pooled whole brains"
 /clone_lib="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 166. 1467
 /codon_start=1
 /product="Similar to phospholipase B"
 /protein_id="AAH42674.1"
 /db_xref="GI:27503749"
 /translation="MPITPTONEPFLRPNSNYTPRIKPAIENWSDFLCEWKRAS
 NSVPSVHOLRPADIVKVAALGSLITTAVARPNMSDLPTRGSGNSIGENIET
 HTLPIILKFNPIYLAGFSTSTWETGRLNVAAGAPARMPAOWDLERKNSPDI
 NLEKDWLVTLFGVNDLCHYCNBPAAHLATEYVQHIQALDILSELPSAFVNVVY
 MELASLYOGGSGKCAIAQNNCTCRHSQSSLEKQIKVMNLIHQGISFSYMYOY
 TOEDPAVVVQFQNTLPLERNGDTDLTFSEDCPHFEDRHAEMVALANMMEP
 VGRKTSNNFTHSRAIKCPSPSPYLYTLIRNSRLIPDQAEAPVLYNAVPAAGVG
 LVVIGITGVVRRRGRRREDPMSLRVAL"

CDS

BASE COUNT 519 a 508 c 437 g 326 t
 ORIGIN

Query Match 30.1%; Score 1319.4; DB 9; Length 1790;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3057 GCTTGAACCACTTGAAGCAAAAAGAGACCTTGACCTGAGAGCAGAGATGCCATCAC 3116
 DB 147 GCTTGAACCACTTGAAGCAAAAAGAGACCTTGACCTGAGAGCAGAGATGCCATCAC 206
 QY 3117 CTGTCCCACTCAGAAATGAGCCCTTCTGAGAACCCCTCGGAATATGTAACAGTACC 3176
 DB 207 CTGTCCCACTCAGAAATGAGCCCTTCTGAGAACCCCTCGGAATATGTAACAGTACC 266
 QY 3177 CATCAAGCCAGCAATTGAGAACTGGGGAGTGACTTCTGTGTATCAGAGTGAAGGCTTC 3236
 DB 267 CATCAAGCCAGCAATTGAGAACTGGGGAGTGACTTCTGTGTATCAGAGTGAAGGCTTC 326
 QY 3237 CAATAGTGTTCACCACTCTGTCCACAGCTCCGACAGCAGACATCAAAATGTTGCGCG 3296
 DB 327 CAATAGTGTTCACCACTCTGTCCACAGCTCCGACAGCAGACATCAAAATGTTGCGCG 386
 QY 3297 CCTGGGTGACTCTCTGACTAGACAGCTGGAGTCCGACCAACCAACTCCAGTACCTACC 3356
 DB 387 CCTGGGTGACTCTCTGACTAGACAGCTGGAGTCCGACCAACCAACTCCAGTACCTACC 446
 QY 3357 CACATCTGGAGGGAGCTCTCTTGAAGCAATTGAGAGGGAGGGAGCAATTGAGAGCTACAC 3416
 DB 447 CACATCTGGAGGGAGCTCTCTTGAAGCAATTGAGAGGGAGGGAGCAATTGAGAGCTACAC 506
 QY 3417 CACACTGCCCAACATTTCTGAAGAGTTCAACCTTACCTCTTGCTTCTCTACAGCAC 3476
 |||

DB 507 CACACTGCCCAACATTTCTGAAGAGTTCAACCTTACCTCTTGCTTCTCTACAGCAC 566
 QY 3477 CTGGAGGGGGAGCAGAGACTAAATGTGGACGGAGGGGCCAGAGCTAGGGACATGCC 3536
 DB 567 CTGGAGGGGGAGCAGAGACTAAATGTGGACGGAGGGGCCAGAGCTAGGGACATGCC 626
 QY 3537 AGCCAGGCGCTGGGACCTGGTAGAGCAATGAAAAAGCCCCAGACATCAACTGGAGAA 3596
 DB 627 AGCCAGGCGCTGGGACCTGGTAGAGCAATGAAAAAGCCCCAGACATCAACTGGAGAA 666
 QY 3597 AGACTGGAAGCTGGTCACTCTTCAATGGGGTCAACAGACTGTGTCTATTACTGTAGAA 3656
 DB 687 AGACTGGAAGCTGGTCACTCTTCAATGGGGTCAACAGACTGTGTCTATTACTGTAGAA 746
 QY 3657 TCCGAGGGCCACTTGGSCACGGAAATGTTCAGCAATCCACAGGCCCTTGACATCTCT 3716
 DB 747 TCCGAGGGCCACTTGGSCACGGAAATGTTCAGCAATCCACAGGCCCTTGACATCTCT 806
 QY 3717 CTCTGAGAGCTTCCCAAGGGCTTTTCTCAACGTGTGTGAGTCACTGAGCTGAGCT 3776
 DB 807 CTCTGAGAGCTTCCCAAGGGCTTTTCTCAACGTGTGTGAGTCACTGAGCTGAGCT 866
 QY 3777 GTACCAAGGCGCAAGCGGGAATGTGCCATGTGCGAGCTCAGAACACTGCACTTGCT 3836
 DB 867 GTACCAAGGCGCAAGCGGGAATGTGCCATGTGCGAGCTCAGAACACTGCACTTGCT 926
 QY 3837 CAGACACTGGCAAGCTCCCTGAGAGCAAGAACTGAAGAACTGGAACCTGCCA 3896
 DB 927 CAGACACTGGCAAGCTCCCTGAGAGCAAGAACTGAAGAACTGGAACCTGCCA 966
 QY 3897 GCATGGCATCTCCAGTTTCTCTCACTGGCAACCAATACACAGCGTGAAGCACTTGCGGT 3956
 DB 987 GCATGGCATCTCCAGTTTCTCTCACTGGCAACCAATACACAGCGTGAAGCACTTGCGGT 1046
 QY 3957 TGTGTGACAGCTTTCTTCCAAAAACACTCACCCCACTGAACGAGAGAGGAGCACTGA 4016
 DB 1047 TGTGTGACAGCTTTCTTCCAAAAACACTCACCCCACTGAACGAGAGAGGAGCACTGA 1106
 QY 4017 CCTCAGCTTCTTCTCCGAGAGCTTTTCACTTCTCAGACCGGGGCAATGCCAGATGCC 4076
 DB 1107 CCTCAGCTTCTTCTCCGAGAGCTTTTCACTTCTCAGACCGGGGCAATGCCAGATGCC 1166
 QY 4077 CATCGCACTCTGGAACAACATGCTGGAACAGTGGGCGGCAAGACTTACCTCCAAACTT 4136
 DB 1167 CATCGCACTCTGGAACAACATGCTGGAACAGTGGGCGGCAAGACTTACCTCCAAACTT 1226
 QY 4137 CACCCAAGCCGAGCAAACTCAAGTCCCTTCTCTGAGAGCCCTTACCTTACACCTT 4196
 DB 1227 CACCCAAGCCGAGCAAACTCAAGTCCCTTCTCTGAGAGCCCTTACCTTACACCTT 1286
 QY 4197 GCGGAAACAGCCGATTTCTCCAGACCAAGGCTGAAGAAAGCCCGAGGTGCTTACTGGGC 4256
 DB 1287 GCGGAAACAGCCGATTTCTCCAGACCAAGGCTGAAGAAAGCCCGAGGTGCTTACTGGGC 1346
 QY 4257 TGTCCAGTGGAGCGGGAGTGGGCTTGTGAGGCAATCATGAGGAGCAAGTGTGGAG 4316
 DB 1347 TGTCCAGTGGAGCGGGAGTGGGCTTGTGAGGCAATCATGAGGAGCAAGTGTGGAG 1406
 QY 4317 GTGCAGAGAGGTGGCGGAGGGAAGATCTCCAAATGAGCTTGCGCATGTGGCCCTTA 4376
 DB 1407 GTGCAGAGAGGTGGCGGAGGGAAGATCTCCAAATGAGCTTGCGCATGTGGCCCTTA 1466
 QY 4377 G 4377
 DB 1467 G 1467
 RESULT 9
 AX713706 2477 bp DNA linear PAT 15-APR-2003
 LOCUS AX713706
 DEFINITION Sequence 390 from Patent Epi293569.
 ACCESSION AX713706
 VERSION AX713706.1 GI:29888564

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Iisogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuno, Y.

TITLE
JOURNAL full-length cDNAs
Patent: EP 1293569-A 390 19-MAR-2003;
Helix Research Institute (Jp) ; Research Association for Biotechnology (Jp)
FEATURES
source location/Qualifiers
1..2477
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 691 a 597 c 591 g 598 t
ORIGIN

Query Match 22.28; Score 970; DB 6; Length 2477;
Best Local Similarity 99.5%; Pred. No. 5,7e-244;
Matches 973; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1279 AGGGTGGCGGAGATGGAACATCGGACCGCTTACCACTCGGGAACATCTCCGGGAA 1338
|||||
6 AGGGTGGCGGAGATGGAACATCGGACCGCTTACCACTCGGGAACATCTCCGGGAA 65
|||||
1339 TTCAACCCCTTCCCTGAGAGGCTTCTCTGTTGGCACTGGGAAAGAACCACTCTAATGCC 1398
|||||
66 TTCAACCCCTTCCCTGAGAGGCTTCTCTGTTGGCACTGGGAAAGAACCACTCTAATGCC 125
|||||
1399 TTCTTAAACAGAGCTGTGGCAGAGGCGGAGCTGAGGATCTACCTGTCCAGGCCAGAGG 1458
|||||
126 TTCTTAAACAGAGCTGTGGCAGAGGCGGAGCTGAGGATCTACCTGTCCAGGCCAGAGG 185
|||||
1459 CTGTGTGACCTGATGGAAGATGACAGAGATACATTTTCAGGAAGACTGGAGATATATA 1518
|||||
186 CTGTGTGACCTGATGGAAGATGACAGAGATACATTTTCAGGAAGACTGGAGATATATA 245
|||||
1519 ACCCTGTTTATAGGCGGCAATGACCTCTGATATTTTGCATGATCTGTCCATATTTCT 1578
|||||
246 ACCCTGTTTATAGGCGGCAATGACCTCTGATATTTTGCATGATCTGTCCATATTTCT 305
|||||
1579 CCCGAGAACTTCAACAGCAACATTTGAAAGGCGCTGAGCATCTCATGCTGAGGTTCT 1638
|||||
306 CCCGAGAACTTCAACAGCAACATTTGAAAGGCGCTGAGCATCTCATGCTGAGGTTCT 365
|||||
1639 CGGGCAATTTGTGAACCTGTGACCGTGTCTTGAAGTCTCTCAACCTGAGAGGAGCTGTACAG 1698
|||||
366 CGGGCAATTTGTGAACCTGTGACCGTGTCTTGAAGTCTCTCAACCTGAGAGGAGCTGTACAG 425
|||||
1699 GAGAAAAAATCTACTACCTCCCAAGGATGATCTCAGGTCTCTGTGCTCCCTGTCTTGAAG 1758
|||||
426 GAGAAAAAATCTACTACCTCCCAAGGATGATCTCAGGTCTCTGTGCTCCCTGTCTTGAAG 485
|||||
1759 TTGTGATGATTAACCAAGAACTTGTACCTCTCATTCGAATTTCAACAGAAATTTACAGAG 1818
|||||
486 TTGTGATGATTAACCAAGAACTTGTACCTCTCATTCGAATTTCAACAGAAATTTACAGAG 545
|||||
1819 AAGAACCAACCACTGATTTGAGAGTGGCGCATATGACACAGGAGAGATTTTACCTGTGTGTT 1878
|||||
546 AAGAACCAACCACTGATTTGAGAGTGGCGCATATGACACAGGAGAGATTTTACCTGTGTGTT 605
|||||
1879 GTGAGCGCGTTCTTTGAAAAAGTGGACATGCAAGGAGGATTTTACCTGTGTGTT 1938
|||||
606 GTGAGCGCGTTCTTTGAAAAAGTGGACATGCAAGGAGGATTTTACCTGTGTGTT 665
|||||
1939 TCTTTCTTGGCTCTGACTGCTTTTCACTTACAGCAGCAAGTCTCACTCCGAGCAGCCAGT 1998
|||||
666 TCTTTCTTGGCTCTGACTGCTTTTCACTTACAGCAGCAAGTCTCACTCCGAGCAGCCAGT 725
|||||

QY 1999 GCTCTGGAACATATGCTGAGCGCTGTGGCCAGAAAGCACTGCTATAAGTTTGA 2058
|||||
DB 726 GCTCTGGAACATATGCTGAGCGCTGTGGCCAGAAAGCACTGCTATAAGTTTGA 785
|||||
QY 2059 AACAGATCAATATCAATATGCTGGAACAGAGTCCAGCCGTTTGTGAGACTTCAAGAAC 2118
|||||
DB 786 AACAGATCAATATCAATATGCTGGAACAGAGTCCAGCCGTTTGTGAGACTTCAAGAAC 845
|||||
QY 2119 AGCATGAGAGGTATGAGACCTGCTGCTCCATGACGAGCAGACCCCTTGTGCTTGCAC 2178
|||||
DB 846 AGCATGAGAGGTATGAGACCTGCTGCTCCATGACGAGCAGACCCCTTGTGCTTGCAC 905
|||||
QY 2179 CCTACCTGAGTGCATGCGCTGAGACCTGAGACATGCAAGTTGTGCTGAGGAGAT 2238
|||||
DB 906 CCTACCTGAGTGCATGCGCTGAGACCTGAGACATGCAAGTTGTGCTGAGGAGAT 965
|||||
QY 2239 TCTCTGACCGCTGCAAT 2256
|||||
DB 966 TCTCTGACCGTAAAGACT 983
|||||

RESULT 10
AK055428 2477 bp mRNA linear PRI 01-AUG-2002
LOCUS Homo sapiens cDNA FLJ30866 fis, clone F088A2004110, highly similar
DEFINITION
ACCESSION AK055428.1 GI:16550150
VERSION AK055428.1 GI:16550150
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Nishi, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, Y., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and Iisogai, T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (base 1 to 2477)
AUTHORS Iisogai, T., Otsuki, T. and Sugiyama, T.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-2001) Takao Iisogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' - 3' end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
source location/Qualifiers
1..2477.
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="F088A2004110"
/tissue_type="brain"
/clone_lib="F088A2"
/dev stage="fetus"
/note="cloning vector: pME18SFL3"
198..1010
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB70920.1"
/db_xref="GI:16550151"

/translation="MKNDTRIHEQEDMKIITLFIGNDLDCDFCNLDVHVSPONFDTNI
GALDILHAEPVPAFVNLVTVLEIVMLRELYOEKRYCPRMILRSPCLVLPEDNST
ELATLIEFNKKFOEKTHOLIESGRVDTREYFVVOPEFMDPMTSPCLSPNSPFA
POCFHSSKSHSRAASALNMNMKEPVGMKTHKFNKINICPNQOVOPFLRTPYKXSM
QGHGTLPCRDRAPSALHPSVHALRPADIQVVALGDSLTVTTLGPQVWVG"
BASE COUNT 691 a 597 c 591 g 598 t
ORIGIN

Query Match 22.2%; Score 970; DB 9; Length 2477;
Best Local Similarity 99.5%; Pred. No. 5.7e-244;
Matches 973; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1279 AGGTCGGCGAGAGTGAACATCGCACCGTTACACCCCTGGCGAATCTCCGAGAA 1338
DB 6 AGGTCGGCGAGAGTGAACATCGCACCGTTACACCCCTGGCGAATCTCCGAGAA 65
QY 1339 TTCAACCTTCTCCGTAAGGGCTTCTGTTGGCACTGGGAAAGAACCACTCTTAATGCC 1398
DB 66 TTCAACCTTCTCCGTAAGGGCTTCTGTTGGCACTGGGAAAGAACCACTCTTAATGCC 125
QY 1399 TTCTTAACGAGCTGTGGAGAGGCGAGCTGAGGATCTACCTGTCCAGGCGAGAGG 1458
DB 126 TTCTTAACGAGCTGTGGAGAGGCGAGCTGAGGATCTACCTGTCCAGGCGAGAGG 185
QY 1459 CTGGTGACCTGATGAAGATGACAGAGGATACCTTTCAGAGAGCTGAAGATTA 1518
DB 186 CTGGTGACCTGATGAAGATGACAGAGGATACCTTTCAGAGAGCTGAAGATTA 245
QY 1519 ACCCTGTTATAGGCGGCAATGACTCTGTGATTTCTGCAATGATCTGTCACTAATCT 1578
DB 246 ACCCTGTTATAGGCGGCAATGACTCTGTGATTTCTGCAATGATCTGTGTCACTAATCT 305
QY 1579 CCCCAAGCTTTCAGACAAATTTGAAAAGCCCTGAGACATCTCCATGCTGAGTTCCT 1638
DB 306 CCCCAAGCTTTCAGACAAATTTGAAAAGCCCTGAGACATCTCCATGCTGAGTTCCT 365
QY 1639 CGGGCAATTTGTGAACCTGTGACGCTGCTTGAATGCTCAACTGAGGAGCTGTACCA 1698
DB 366 CGGGCAATTTGTGAACCTGTGACGCTGCTTGAATGCTCAACTGAGGAGCTGTACCA 425
QY 1699 GAGAAAAAGTCTACCTGCCAAGATGATCTCTGTGTCTGTCTGTCTGTGAAG 1758
DB 426 GAGAAAAAGTCTACCTGCCAAGATGATCTCTGTGTCTGTCTGTCTGTGAAG 485
QY 1759 TTGTGATTAATCAACAACTTGTACCTCATGGAATTCACAAAGATTTTCAGAG 1818
DB 486 TTGTGATTAATCAACAACTTGTGTACCTCATGGAATTCACAAAGATTTTCAGAG 545
QY 1819 AAGACCCCAACTGATTTGAGAGTGGGCGATATGACAAAGGAGATTTTACTGTGATT 1878
DB 546 AAGACCCCAACTGATTTGAGAGTGGGCGATATGACAAAGGAGATTTTACTGTGATT 605
QY 1879 GTGACGCGCTTCTTGAAGAACTGTGACATGCCAAAGACCTCGAGAGATTTCTGCAAC 1938
DB 606 GTGACGCGCTTCTTGAAGAACTGTGACATGCCAAAGACCTCGAGAGATTTCTGCAAC 665
QY 1939 TCTTTCTTGGCTCTGACTGTTCCTCACTTGAAGAGATCTCACTCCGAGAGCAAGT 1998
DB 666 TCTTTCTTGGCTCTGACTGTTCCTCACTTGAAGAGATCTCACTCCGAGAGCAAGT 725
QY 1999 GCTCTGTGAAGAACTATGCTGAGCGCTGTGGCGAGAGCACTGTGCTAATGATTGAA 2058
DB 726 GCTCTGTGAAGAACTATGCTGAGCGCTGTGGCGAGAGCACTGTGCTAATGATTGAA 785
QY 2059 AACCAAGATCAATATCAATGTCCGAAACAGAGTCAGCGCTTCTGAGAGCTTCAAGAAC 2118
DB 786 AACCAAGATCAATATCAATGTCCGAAACAGAGTCAGCGCTTCTGAGAGCTTCAAGAAC 845
QY 2119 AGCATGAGAGGTATGAGAGCTTGGCTGCTGACAGGAGAGAGCCCTTCTGCTTGAC 2178
DB 846 AGCATGAGAGGTATGAGAGCTTGGCTGCTGACAGGAGAGAGCCCTTCTGCTTGAC 905
QY 2179 CTAACCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2238

DB 906 CTAACCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 965
QY 2239 TCTGTGACGGTGGCAAT 2256
DB 966 TCTGTGACGGTGGCAAT 983

RESULT 11
AX417821
LOCUS AX417821 608 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 5 from Patent WO0231161.
ACCESSION AX417821
VERSION AX417821.1 GI:21522939
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 Zhu Z.
Regulation of human phospholipase-1-like enzyme
Patent: WO 0231161-A 5 18-APR-2002;
BAYER AG (DE)
FEATURES
source location/Qualifiers
1..608
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 140 a 187 c 158 g 123 t
ORIGIN

Query Match 13.2%; Score 576.2; DB 6; Length 608;
Best Local Similarity 97.8%; Pred. No. 3.2e-140;
Matches 584; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2654 TGAACGCTCTGATAGAGAGTGGCCAGAGCTCTGATCAACTGCTGATCTTCTGATCC 2713
DB 11 TGAACGCTCTGATAGAGAGTGGCCAGAGCTCTGATCAACTGCTGATCTTCTGATCC 70
QY 2714 CCACTATCATGCGGAGGCTTCTTGGGAAACCCAGACAAAGTGCAGAGGAGGCA 2773
DB 71 CCACTATCATGCGGAGGCTTCTTGGGAAACCCAGACAAAGTGCAGAGGAGGCA 130
QY 2774 GCGTTTGTGTATCTGCTTCTGACCTTCCGAGAACTCCCAAGCTTACAGGCTG 2833
DB 131 GCGTTTGTGTATCTGCTTCTGACCTTCCGAGAACTCCCAAGCTTACAGGCTG 190
QY 2834 AGGCTTCAAGCGAGGCTTCTGAGAGAGAGTGGAGGAGTGGAGGCTGAGGCGCTATG 2893
DB 191 AGGCTTCAAGCGAGGCTTCTGAGAGAGAGTGGAGGAGTGGAGGCTGAGGCGCTATG 250
QY 2894 ACAAGCAGAGAGACTTCTGTGTGCTGCTGACAGCCCTTCTTCCAGAACTACAGCTCTG 2953
DB 251 ACAAGCAGAGAGACTTCTGTGTGCTGCTGACAGCCCTTCTTCCAGAACTACAGCTCTG 310
QY 2954 TCTGTGCGAGTGGCTCCAGATAGCTCTTCTTGGCCCAAGCTGATCCACCAATC 3013
DB 311 TCTGTGCGAGTGGCTCCAGATAGCTCTTCTTGGCCCAAGCTGATCCACCAATC 370
QY 3014 AGAAATTCACCTCCAGCTGCGGAGAGCCCTTGGAGCAATGCTTGAACCACTTGGAA 3073
DB 371 AGAAATTCACCTCCAGCTGCGGAGAGCCCTTGGAGCAATGCTTGAACCACTTGGAA 430
QY 3074 GCAAAACAGAGACCTTGAAGCTGAGAGAGAGATGCCATCACTGTCCCACTCAGAAATG 3133
DB 431 GCAAAACAGAGACCTTGAAGCTGAGAGAGAGATGCCATCACTGTCCCACTCAGAAATG 490
QY 3134 AGCCCTTCTGAGAAACCTCGAATAGTAACTACAGTACCCCATCAAGCCAGCATTTG 3193
DB 491 AGCCCTTCTGAGAAACCTCGAATAGTAACTACAGTACCCCATCAAGCCAGCATTTG 550
QY 3194 ABAAGTGGGAGTACCTTCTGTGTACAGAGTGGAGGCTTCAATATGTTCCA 3250

```

Db      551 AGAAGTGGGAGTGAAGTCTTCTGTGTGACAGAGTGAAGTCTCAATAGTGTCCAA 607
RESULT 12
LOCUS   AX417820/c
DEFINITION Sequence 4 from Patent WO0231161.
ACCESSION AX417820
VERSION  AX417820.1 GI:21522938
KEYWORDS
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Zhu, Z.
TITLE    Regulation of human phospholipase-1-like enzyme
JOURNAL  Patent: WO 0231161-A 4 18-Apr-2002;
          BAYER AG (DE)
FEATURES
        source      1..583
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
BASE COUNT  114 a 155 c 190 g 124 t
ORIGIN
Query Match      8.5%; Score 372.2; DB 6; Length 583;
Best Local Similarity 99.2%; Pred. No. 1.9e-86;
Matches 374; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  4001 AGAGAGGGGACACTGACCTTCTTCCCGAGACCTTTTCACTTCTCAGACCGCG 4060
Db  583 AGAGAGGGGACACTGACCTTCTTCCCGAGACCTTTTCACTTCTCAGACCGCG 524
QY  4061 GGCATGCCGAGATGCGCATGCACTCTGAAACAACATCTGAAACAGTGGCGCGAAGA 4120
Db  523 GGCATGCCGAGATGCGCATGCACTCTGAAACAACATCTGAAACAGTGGCGCGAAGA 464
QY  4121 CTACCTTCAACAACCTTCAACCCAGCGGAGCAAACTCAAGTGCCTCTCTGAGAGCC 4180
Db  463 CTACCTTCAACAACCTTCAACCCAGCGGAGCAAACTCAAGTGCCTCTCTGAGAGCC 404
QY  4181 CTACCTTCAACCCAGCGGAGCAAACTCAAGTGCCTCTCTGAGAGCC 4240
Db  403 CTACCTTCAACCCAGCGGAGCAAACTCAAGTGCCTCTCTGAGAGCC 344
QY  4241 AGGTGCTCTACTGAGGTGAGAGAGTGGCGGAGTGGCGCTTGTGGGATCATCG 4300
Db  343 AGGTGCTCTACTGAGGTGAGAGAGTGGCGGAGTGGCGCTTGTGGGATCATCG 284
QY  4301 GGAAGTGTCTGAGAGTGAAGAGAGTGGCGGAGAGATCTTCAATGAGCTTGC 4360
Db  283 GGAAGTGTCTGAGAGTGAAGAGAGTGGCGGAGAGATCTTCAATGAGCTTGC 224
QY  4361 GCACGTGGCGCTCTTAG 4377
Db  223 GCACGTGGCGCTCTTAG 207

```

```

REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
          Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
          Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
          Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
          Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
          Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
          Scheetz, T.E., Brownstein, M.J., Umedin, T.B., Toshiyuki, S.,
          Carninci, P., Prange, C., Rahe, S.S., Loquellano, N.A., Peters, G.J.,
          Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
          McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
          Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hult, J., Gibbs, R.A.,
          Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
          Fahey, J., Helton, E., Ketterman, M., Madan, A., Rodriguez, S.,
          Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shchepenko, Y.,
          Bouffard, G.G., Blakeley, R.W., Touchman, D.W., Green, E.D.,
          Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
          Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E.,
          Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
          Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
          Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
          22388257
JOURNAL MEDLINE
PUBMED 12477932
REFERENCE 2 (bases 1 to 820)
AUTHORS Strausberg, R.
TITLE    Direct Submission
JOURNAL Submitted (02-JUL-2002) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK
COMMENT NIH-MGC Project URL: http://mgc.ncl.nih.gov
          Contact: MGC help desk
          Email: cgapob-re@mail.nih.gov
          Tissue Procurement: Gilbert Smith, Ph.D.
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
          DNA Sequencing by: National Institutes of Health Intramural
          Sequencing Center (NISC),
          Gaithersburg, Maryland;
          Web site: http://www.nisc.nih.gov/
          Contact: nisc.mgc@hgtl.nih.gov
          Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
          Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
          Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
          Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaapi, R.,
          Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
          McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W.,
          Turegion, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
          Young, A., Zhang, L.-H. and Green, E.D.
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov
          Series: IRAC Plate: 65 Row: 1 Column: 10
          This clone was selected for full length sequencing because it
          passed the following selection criteria: Hexamer frequency ORF
          analysis.
FEATURES
        source      1..820
                    /organism="Mus musculus"
                    /mol_type="mRNA"
                    /strain="FVB/N-3"
                    /db_xref="taxon:10090"
                    /clone="MGC40917 IMAGE:5373588"
                    /tissue_type="Mammary tumor, MMTV-LTR/INT3 model, 5 month
                    old mouse. Taken by biopsy."
                    /clone_1ib="NCI CGAP_Mam2"
                    /lab_host="DH10B"
                    /note="Vector: pCMV-SPORT6"
        gene        1..820
                    /gene="MGC40917"
                    /db_xref="LocusID:231108"
        CDS          94..654

```

```

/codon_start=1
/product="hypothetical protein MG40917"
/protein_id="AAH33606.1"
/db_xref="GI:21706930"
/db_xref="LOCUSID:231108"
/translation="MOBKLTANNLOSISELSYHRYMEREDPAVTVDPFFNTIP
LNREGLDITFESDCEYFSDRGHAELAIALNMNMLEPVMKTSNNFIYNRKICP
SPERPPLYLIRNSQLPDKAEPBNALYMAVPAVAAIGLAVGILGVTLMRTVPOOE
EEEDPTLPNTSVTVQDAVSEKRLKAGN"
BASE COUNT      266 a      197 c      188 g      169 t
ORIGIN

```

```

Query Match      7.4%; Score 322.2; DB 10; Length 820;
Best Local Similarity 75.1%; Pred. No. 3e-73;
Matches 402; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

```

```

QY      3809 TGGCAGCTCAAACTGCACTTGGCTTCAGACCTGCGAAAGTCCCTGGAGAGCAAG 3868
DB      41 TGTCTGTAGGAAAACTGCAAGTTGCTTAAGACCTCCAAAACCTTCACAGCATGACAG 100
QY      3869 AACTGAAGAAAGTGAAGTGGACCTCCAGATGGCATCTCAGTTTCTCTACTGGCACC 3928
DB      101 AGCTGAAGAACTAACTGGAACCTCCAGAGCGCATCTCGAGCTTCTCTATTGGCACC 160
QY      3929 AATACA CACAGCGTGAAGACTTTCGGGTGTGTGTCAGACCTTTCTTCCAAAACACTCA 3988
DB      161 GGTACATGAGAGCGTGAAGACTTTCGGGTGTGTGTCAGACCTTTCTTCCAAAACACTCA 220
QY      3989 CCCCAGTGAACGAGAGAGGAGCACTGACCTCAGCTTCTTCTCCAGAGACTGTTTCACT 4048
DB      221 TCCCACTGAATGAGCGTGAAGCGCTGGAACCTCACTTCTTCTGGAAGACTGTTTCACT 280
QY      4049 TCTCAGACGCGGGGATGCGGAGATGGGCATGCGACCTGGAACCAATCTGGAACACAG 4108
DB      281 TCTCAAGCGGTGGCATGTGAGATGGCCATGTGCCCTCTGAATATCATCTGGAACACAG 340
QY      4109 TGGGCGCGAAGACTTCTCCAACTTCAACCAAGCCGAGCCAAACTCAAGTGCCT 4168
DB      341 TGGGCGTGAAGAACTCTCCAACTTCAATATATCAACCAACCAACTCAAGTGCCT 400
QY      4169 CTCTGAGAGCCCTTACCTCTTACACCTGCGGAAACAGCCGATGCTCCAGACCAAGCTG 4228
DB      401 CACTGAAAGGCTTCTCTTACACCTCCGGAATGTGACGTTCTTCCAGACCAAGCTG 460
QY      4229 AAGAACCCCGGAGGTCTCTACTGCGGTCTCCAGTGGAGGAGGAGTGGGCTTGTG 4288
DB      461 AAGAACCCCTCAATGCACTTACTGCGGAGTGCAGAGTGCAGATATGATGCGCTGCGAG 520
QY      4289 TGGGCATCATCGGAGACAGTGTCTGAGAGTGCAGAGAGGTGGCCGAGGAGAGA 4343
DB      521 TTGGCATCTTGGAGATGATGTTGTGAGAACTGTGAAACCCGTCCAACAGAGAGA 575

```

```

RESULT 14
AX417822      265 bp      DNA      linear      PAT 18-JUN-2002.
LOCUS      AX417822      Sequence 6 from Patent W00231161.
DEFINITION      AX417822
ACCESSION      AX417822
VERSION      AX417822.1 GI:21522940
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
                Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1. Zhu, Z.
  Regulation of human phospholipase-1 like enzyme.
  JOURNAL      BAYER AG (DE)
                Location/Qualifiers
FEATURES
  source
    1..265
    /organism="Homo sapiens"
    /mol_type="genomic DNA"

```

```

BASE COUNT      73 a      80 c      62 g      50 t
ORIGIN

```

```

Query Match      5.7%; Score 251; DB 6; Length 265;
Best Local Similarity 100.0%; Pred. No. 1.6e-54;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3042 CCTTGGACCAATATGCTTGAACCACTGGAACCAAAACAGAGCCCTGAGCTGAGAGC 3101
DB      15 CCTTGGACCAATATGCTTGAACCACTGGAACCAAAACAGAGCCCTGAGCTGAGAGC 74
QY      3102 AGAGATGCCCATCACTGTCTCCACTGAGATGAGCCCTTCTGAGAACCCCTCGAATAG 3161
DB      75 AGAGATGCCCATCACTGTCTCCACTGAGATGAGCCCTTCTGAGAACCCCTCGAATAG 134
QY      3162 TAACTACAGTACCCCATCAAGCCAGCCATTTGAGAACTGGGGGAGTACTTCTGTGTAC 3221
DB      135 TAACTACAGTACCCCATCAAGCCAGCCATTTGAGAACTGGGGGAGTACTTCTGTGTAC 194
QY      3222 AGAGTGAAGGCTTCCAAATGATGTTCCAACTGTCTCCACAGCTCCGACAGAGACAT 3281
DB      195 AGAGTGAAGGCTTCCAAATGATGTTCCAACTGTCTCCACAGCTCCGACAGAGACAT 254
QY      3282 CAAAGTGTGTG 3292
DB      255 CAAAGTGTGTG 265

```

```

RESULT 15
AX417823      446 bp      DNA      linear      PAT 18-JUN-2002
LOCUS      AX417823/c
DEFINITION      Sequence 7 from Patent W00231161.
ACCESSION      AX417823
VERSION      AX417823.1 GI:21522941
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
                Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1. Zhu, Z.
  Regulation of human phospholipase-1 like enzyme
  JOURNAL      BAYER AG (DE)
                Location/Qualifiers
FEATURES
  source
    1..446
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"

```

```

BASE COUNT      86 a      122 c      147 g      90 t      1 others
ORIGIN

```

```

Query Match      5.4%; Score 237.4; DB 6; Length 446;
Best Local Similarity 99.2%; Pred. No. 6.4e-51;
Matches 238; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      4138 ACCACAGCCGAGCCAACTCAAGTCCCTCTCTGAGAGCCCTTACCTTACACCTTG 4197
DB      446 ACCACAGCCGAGCCAACTCAAGTCCCTCTCTGAGAGCCCTTACCTTACACCTTG 387
QY      4198 CGGAACAGCCGATTTGCTCCAGACCAAGGCTGAGAAAGCCCCCGAGGTGCTTACTGGGCT 4257
DB      386 CGGAACAGCCGATTTGCTCCAGACCAAGGCTGAGAAAGCCCCCGAGGTGCTTACTGGGCT 327
QY      4258 GTCCAGTGCAGCGGAGAGTGGCTTGTGTGTGGCATCATCGGAGACAGTGTCTGGAGG 4317
DB      326 GTCCAGTGCAGCGGAGAGTGGCTTGTGTGTGGCATCATCGGAGACAGTGTCTGGAGG 267
QY      4318 TGCAGAGAGTGTGCGGAGGAGAAATCTTCAATAGACCTTGCGCACTGTGCGCTCTTAG 4377
DB      266 TGCAGAGAGTGTGCGGAGGAGAAATCTTCAATAGACCTTGCGCACTGTGCGCTCTTAG 207

```

Wed Jan 7 10:03:20 2004

us-10-054-691-1.rge

Page 22

Search completed: January 6, 2004, 12:19:30
Job time : 15568 secs

```

US-10-054-691-1
Sequence 1, Application US/10054691
Publication No. US20020115846A1
GENERAL INFORMATION:
APPLICANT: Yu, Xuanchuan
APPLICANT: Miranda, Maricar
TITLE OF INVENTION: No. US20020115846A1 Human Lipase and Polynucleotides Encoding th
FILE REFERENCE: LEX-0303-USA
CURRENT APPLICATION NUMBER: US/10/054,691
CURRENT FILING DATE: 2001-01-22
PRIORITY APPLICATION NUMBER: US 60/264,049
PRIORITY FILING DATE: 2001-01-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 4377
TYPE: DNA
ORGANISM: homo sapiens
US-10-054-691-1

Query Match      100.0%; Score 4377; DB 14; Length 4377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGGGGCTGCGCGCGAGGATTTTCTCCTGGAAGCTGCTGCTTCTGGGGCAAGGACC 60
Db      1 ATGGGGCTGCGCGCGAGGATTTTCTCCTGGAAGCTGCTGCTTCTGGGGCAAGGACC 60

QY      61 CCTAGATCCATCTCTCTCTTGAAGAAGTAACTTGAAGGAGGAGCTATGCGCAGAGACC 120
Db      61 CCTAGATCCATCTCTCTCTTGAAGAAGTAACTTGAAGGAGGAGCTATGCGCAGAGACC 120

QY      121 CTGAAGATTCCTCATCTTCCCATGCAACCAATTAATTGAGAGTAATGCTTTCTTAA 180
Db      121 CTGAAGATTCCTCATCTTCCCATGCAACCAATTAATTGAGAGTAATGCTTTCTTAA 180

RESULT 1
US-10-054-691-1
Sequence 1, Application US/10054691
Publication No. US20020115846A1
GENERAL INFORMATION:
APPLICANT: Yu, Xuanchuan
APPLICANT: Miranda, Maricar
TITLE OF INVENTION: No. US20020115846A1 Human Lipase and Polynucleotides Encoding th
FILE REFERENCE: LEX-0303-USA
CURRENT APPLICATION NUMBER: US/10/054,691
CURRENT FILING DATE: 2001-01-22
PRIORITY APPLICATION NUMBER: US 60/264,049
PRIORITY FILING DATE: 2001-01-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 4377
TYPE: DNA
ORGANISM: homo sapiens
US-10-054-691-1

ALIGNMENTS

c 16 41.8 1.0 11056 11 US-09-974-899-98 Sequence 98, Appl
c 17 41.8 1.0 31058 10 US-09-764-281-6138 Sequence 6138, Ap
c 18 41.8 1.0 169998 13 US-10-380-931-24 Sequence 24, Appl
c 19 41.8 1.0 197496 10 US-09-877-177-10 Sequence 10, Appl
c 20 41.6 1.0 505 11 US-09-918-999-31558 Sequence 31558, A
c 21 41 0.9 653 15 US-10-184-644-402 Sequence 402, App
c 22 41 0.9 653 15 US-10-184-644-402 Sequence 402, App
c 23 40.8 0.9 4689 10 US-09-861-289-34 Sequence 34, Appl
c 24 40.8 0.9 4689 10 US-09-861-289-34 Sequence 34, Appl
c 25 40.8 0.9 4689 11 US-09-988-384B-34 Sequence 34, Appl
c 26 40.8 0.9 4689 11 US-09-836-821-34 Sequence 34, Appl
c 27 40.8 0.9 4689 13 US-10-271-882-34 Sequence 34, Appl
c 28 40.8 0.9 36778 10 US-09-861-289-5 Sequence 5, Appli
c 29 40.8 0.9 36778 10 US-09-861-289-5 Sequence 5, Appli
c 30 40.8 0.9 36778 11 US-09-836-821-5 Sequence 5, Appli
c 31 40.8 0.9 36778 11 US-10-271-882-48 Sequence 48, Appl
c 32 40.8 0.9 37948 11 US-09-988-384B-5 Sequence 5, Appli
c 33 40.8 0.9 38506 11 US-09-793-708-19 Sequence 19, Appl
c 34 40.8 0.9 38506 13 US-10-201-365-1 Sequence 1, Appli
c 35 40.8 0.9 38506 13 US-10-160-538-19 Sequence 19, Appl
c 36 40.2 0.9 877 12 US-10-369-493-30712 Sequence 30712, A
c 37 40.2 0.9 1308 15 US-10-156-761-6349 Sequence 6349, Ap
c 38 40.2 0.9 9025608 15 US-10-181-779-16 Sequence 16, Appl
c 39 38.8 0.9 1896 13 US-10-175-523-57 Sequence 57, Appl
c 40 38.8 0.9 228883 15 US-10-156-761-2812 Sequence 2812, Ap
c 41 38.6 0.9 1428 15 US-10-184-644-312 Sequence 312, App
c 42 38 0.9 802 15 US-10-184-634-312 Sequence 312, App
c 43 38 0.9 802 13 US-09-908-975-228 Sequence 228, App
c 44 37.8 0.9 65 13 US-10-140-472-412 Sequence 412, App
c 45 37.8 0.9 1184 13 US-10-140-472-412 Sequence 412, App

```

181 TCAGTCACTCTCTGAGCCCTTCTGATATTAATTTGTGCGAGCCATTGCGAATCTGGAA 240
181 TCAGTCACTCTCTGAGCCCTTCTGATATTAATTTGTGCGAGCCATTGCGAATCTGGAA 240
241 ATTCTCTCAAGCCAGGAGCGGGGATCTGAGAGAACAGACTGAGCTGAAAGGCCAG 300
241 ATTCTCTCAAGCCAGGAGCGGGGATCTGAGAGAACAGACTGAGCTGAAAGGCCAG 300
301 CAGGTGCGATGGGAGGATGACAGCTCTTTCAGACATCATCATGATTTCAAGCTTCT 360
301 CAGGTGCGATGGGAGGATGACAGCTCTTTCAGACATCATCATGATTTCAAGCTTCT 360
361 GTTCCATGCTGCTGTGTCACACTGAGAAAGAGTCTATCCCAAGATGAGTGTCTGAGAC 420
361 GTTCCATGCTGCTGTGTCACACTGAGAAAGAGTCTATCCCAAGATGAGTGTCTGAGAC 420
421 TTGTGATTCAGGCTCAAGAACTGCTGAGAAACATGAAAGAGAACTTGCACCTTT 480
421 TTGTGATTCAGGCTCAAGAACTGCTGAGAAACATGAAAGAGAACTTGCACCTTT 480
481 CAATTTGACTGGAAGCTCATCAATGTTCTTCTGATGATGCAAGCCAGTGTACCTGTC 540
481 CAATTTGACTGGAAGCTCATCAATGTTCTTCTGATGATGCAAGCCAGTGTACCTGTC 540
541 CCTCTGCTCAACAGAAATGGGCTTTCGCGCGGCGCGCTGATGAGCTGATGGGGTCTG 600
541 CCTCTGCTCAACAGAAATGGGCTTTCGCGCGGCGCGCTGATGAGCTGATGGGGTCTG 600
601 GACTAATCTGAGAGAGAGTCTCCAGAGCATTTGTAACTGTTGAGCTCTCTGAGGTT 660
601 GACTAATCTGAGAGAGAGTCTCCAGAGCATTTGTAACTGTTGAGCTCTCTGAGGTT 660
601 GACTAATCTGAGAGAGAGTCTCCAGAGCATTTGTAACTGTTGAGCTCTCTGAGGTT 660
601 GACTAATCTGAGAGAGAGTCTCCAGAGCATTTGTAACTGTTGAGCTCTCTGAGGTT 660
661 GCGAGGCTCTGCTGCTGATCAAGGCACTTGTGCTGAGCCCTGACAGAGCCCTGTAAT 720
661 GCGAGGCTCTGCTGCTGATCAAGGCACTTGTGCTGAGCCCTGACAGAGCCCTGTAAT 720
721 TGCTCAGAGAGAGACACCCGCTGCGCAAGGTGCTGATGCAAGTGTCTTATCAGAAAGCC 780
721 TGCTCAGAGAGAGACACCCGCTGCGCAAGGTGCTGATGCAAGTGTCTTATCAGAAAGCC 780
781 TGCTCAGAGAGAGACACCCGCTGCGCAAGGTGCTGATGCAAGTGTCTTATCAGAAAGCC 780
781 TGCTCAGAGAGAGACACCCGCTGCGCAAGGTGCTGATGCAAGTGTCTTATCAGAAAGCC 780
841 CAGGCTTCTTCTGATGAGACACCCCACTCTACACTGAGAGAGCCCTGCACTCAGAGAT 900
841 CAGGCTTCTTCTGATGAGACACCCCACTCTACACTGAGAGAGCCCTGCACTCAGAGAT 900
841 CAGGCTTCTTCTGATGAGACACCCCACTCTACACTGAGAGAGCCCTGCACTCAGAGAT 900
841 CAGGCTTCTTCTGATGAGACACCCCACTCTACACTGAGAGAGCCCTGCACTCAGAGAT 900
901 TCTACACAGCTGCTGCTGAGATCTCTGAAATGAGATGATGAGAGAGAGAGAGAT 960
901 TCTACACAGCTGCTGCTGAGATCTCTGAAATGAGATGATGAGAGAGAGAGAGAT 960
961 GAGCATTGAGTGTAAACACGAGAGGCAATGAAAGTGTCTCTCAGAGAGAGCCCTAT 1020
961 GAGCATTGAGTGTAAACACGAGAGGCAATGAAAGTGTCTCTCAGAGAGAGCCCTAT 1020
1021 CTGTTCAGCTACAGAAACAGCACTACCTGACAGACTGAGAGAGAGAGAGAGAT 1080
1021 CTGTTCAGCTACAGAAACAGCACTACCTGACAGACTGAGAGAGAGAGAGAGAT 1080
1081 GAGGTAGAT 1140
1081 GAGGTAGAT 1140
1141 ACCTCAGTTCATGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1200
1141 ACCTCAGTTCATGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1200
1201 CTACAGGAT 1260
1201 CTACAGGAT 1260
1261 TACGAT 1320
1261 TACGAT 1320

1261 TACGAT 1320
1321 GCGAATATCTCCGAGAAATTCACACCTTCCCTGAAAGGCTTCTCTGAGCACTGGGAAA 1380
1321 GCGAATATCTCCGAGAAATTCACACCTTCCCTGAAAGGCTTCTCTGAGCACTGGGAAA 1380
1381 GAAACAGTCTTATGCTTCTTAAACAGAGCTGTGAGAGAGAGAGAGAGAGAT 1440
1381 GAAACAGTCTTATGCTTCTTAAACAGAGCTGTGAGAGAGAGAGAGAGAGAT 1440
1441 CCTGTCAGGCGCAGAGAGCTGTGAGAGCTGATGAGAGAGAGAGAGAGAT 1500
1441 CCTGTCAGGCGCAGAGAGCTGTGAGAGCTGATGAGAGAGAGAGAGAGAT 1500
1501 GAGAGCTGAGAGATTAATTAATCCCTGTTATAGGGGAGAGAGAGAGAGAT 1560
1501 GAGAGCTGAGAGATTAATTAATCCCTGTTATAGGGGAGAGAGAGAGAGAT 1560
1561 GATCTGATCCATATCTCTCCAGAACTTCAAGAGACATTTGAAAGGCTTGAATC 1620
1561 GATCTGATCCATATCTCTCCAGAACTTCAAGAGACATTTGAAAGGCTTGAATC 1620
1621 CTCCATGCTGAGAGGCTTCTCCGAGCATTTGAACTGTGACCTGCTGAGATCTCAAC 1680
1621 CTCCATGCTGAGAGGCTTCTCCGAGCATTTGAACTGTGACCTGCTGAGATCTCAAC 1680
1681 CTGAGAGAGAGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1740
1681 CTGAGGAGAGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1740
1741 TGTCCTGCTGCTGAGAGATTTGATGATTAATCAACAGAACTTGTATCTGATC 1800
1741 TGTCCTGCTGCTGAGAGATTTGATGATTAATCAACAGAACTTGTATCTGATC 1800
1801 AACAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1860
1801 AACAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1860
1861 GAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1920
1861 GAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1920
1921 GAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1980
1921 GAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1980
1981 GAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2040
1981 GAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2040
2041 ACTGCTATTAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2100
2041 ACTGCTATTAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2100
2101 CTGAT 2160
2101 CTGAT 2160
2161 GCGCCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2220
2161 GCGCCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2220
2221 GTGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2280
2221 GTGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2280
2281 CTGAT 2340
2281 CTGAT 2340
2341 CTGAT 2400
2341 CTGAT 2400

Db 2341 CTGAGAAATGTGACCACTTACTATATCTTCCGGAGTTTAAACAGAAACCTCAAGGC 2400
 Qy 2401 TAGCCGTGGGACGGGTGATGCAATGACAGCAATGATCTTCTCAATCAAGCTGTTCCC 2460
 Db 2401 TAGCCGTGGGACGGGTGATGCAATGACAGCAATGATCTTCTCAATCAAGCTGTTCCC 2460
 Qy 2461 GGAGCAAAAGCTGAGATCTTATGAGCCAAAGTCCAACTCTGATGCAAGATGAAAGAT 2520
 Db 2461 GGAGCAAAAGCTGAGATCTTATGAGCCAAAGTCCAACTCTGATGCAAGATGAAAGAT 2520
 Qy 2521 GATCATAGATTAATTTCCATGAAAGATGGAAGTCTCAAGTCTGATCGAGAGC 2580
 Db 2521 GATCATAGATTAATTTCCATGAAAGATGGAAGTCTCAAGTCTGATCGAGAGC 2580
 Qy 2581 GATTATGATGATGATGCAAGATGGAATGCTGATCTGACCAACTTGTGATCCAT 2640
 Db 2581 GATTATGATGATGATGCAAGATGGAATGCTGATCTGACCAACTTGTGATCCAT 2640
 Qy 2641 CTCGGCAATGCTTGGACGCTCTGCAATGAGAGTCCAGAGTCTGTGCACTCTG 2700
 Db 2641 CTCGGCAATGCTTGGACGCTCTGCAATGAGAGTCCAGAGTCTGTGCACTCTG 2700
 Qy 2701 GACTTCTGTAACCCCACTATCATGCGGAGGTCTTCTGGGAAACCCAGCAAGTCCCA 2760
 Db 2701 GACTTCTGTAACCCCACTATCATGCGGAGGTCTTCTGGGAAACCCAGCAAGTCCCA 2760
 Qy 2761 GTGACAGAGGCGAGCGTTTGTGTAATGCGGTCTGACCTTGGGGAGAACTCCCAAG 2820
 Db 2761 GTGACAGAGGCGAGCGTTTGTGTAATGCGGTCTGACCTTGGGGAGAACTCCCAAG 2820
 Qy 2821 CTAGCCAGGCTGAGAGGCTTCAAGCCAGCTTACCGAGAGCAATGCCAGAGCTGTGG 2880
 Db 2821 CTAGCCAGGCTGAGAGGCTTCAAGCCAGCTTACCGAGAGCAATGCCAGAGCTGTGG 2880
 Qy 2881 TCAGGCGGCTATGACACGACAGAGAGCTTCTGTGTGCTGAGGCCCTTCTTCAAG 2940
 Db 2881 TCAGGCGGCTATGACACGACAGAGAGCTTCTGTGTGCTGAGGCCCTTCTTCAAG 2940
 Qy 2941 ATCCAGTCCCTGTCCTGCGGAGATGAGGCTCCAGATACGTCCTTGTGCGCCAGACTG 3000
 Db 2941 ATCCAGTCCCTGTCCTGCGGAGATGAGGCTCCAGATACGTCCTTGTGCGCCAGACTG 3000
 Qy 3001 ATCCAGTCCCTGTCCTGCGGAGATGAGGCTCCAGATACGTCCTTGTGCGCCAGACTG 3060
 Db 3001 ATCCAGTCCCTGTCCTGCGGAGATGAGGCTCCAGATACGTCCTTGTGCGCCAGACTG 3060
 Qy 3061 GAAACAATTGGAAGCAAAACAGAGCCCTGGAACCTGAGACAGAGATGCCATCACTGT 3120
 Db 3061 GAAACAATTGGAAGCAAAACAGAGCCCTGGAACCTGGAACCTGAGACAGAGATGCCATCACTGT 3120
 Qy 3121 CCCACTCAGAAATAGCCCTTCTGAGAAACCTTCCGAAATGTAATCAAGATACCCATC 3180
 Db 3121 CCCACTCAGAAATAGCCCTTCTGAGAAACCTTCCGAAATGTAATCAAGATACCCATC 3180
 Qy 3181 AAGCCAGCAATTGGAAGCTGAGGAGTGAATCTTCTGTGTAAGAGTGAAGGCTTCAAT 3240
 Db 3181 AAGCCAGCAATTGGAAGCTGAGGAGTGAATCTTCTGTGTAAGAGTGAAGGCTTCAAT 3240
 Qy 3241 AGTGTTCACACCTTGTCCACAGCTCCGACAGACAGATCAAAAGTGTGCGCTG 3300
 Db 3241 AGTGTTCACACCTTGTCCACAGCTCCGACAGACAGATCAAAAGTGTGCGCTG 3300
 Qy 3301 GGTGACTCTTGTACTACAGAGTGGAGTGCACAAACATCTCAATGATCAACCCACA 3360
 Db 3301 GGTGACTCTTGTACTACAGAGTGGAGTGCACAAACATCTCAATGATCAACCCACA 3360
 Qy 3361 TCTTGGAGGGAATCTTGTGAGATTTGAGGGAATGGAATCTTGAATCTTCAACACA 3420
 Db 3361 TCTTGGAGGGAATCTTGTGAGATTTGAGGGAATGGAATCTTGAATCTTCAACACA 3420
 Qy 3421 CTGCCCCAATTTGTGAAGAGTTCAACCTTACTCTTGTGCTTCTTACAGACACTG 3480
 Db 3421 CTGCCCCAATTTGTGAAGAGTTCAACCTTACTCTTGTGCTTCTTACAGACACTG 3480

Qy 3481 GAGGGGACAGACGATTAATGTGCGAGCGAAAGGGGCGAGACTAGGGAATGCCAGCC 3540
 Db 3481 GAGGGGACAGACGATTAATGTGCGAGCGAAAGGGGCGAGACTAGGGAATGCCAGCC 3540
 Qy 3541 CAGGCTGAGGACTGTGTAAGCGAATGAAAAACGCCCCGACATCAACTGAGAAAGAC 3600
 Db 3541 CAGGCTGAGGACTGTGTAAGCGAATGAAAAACGCCCCGACATCAACTGAGAAAGAC 3600
 Qy 3601 TGGAACTGTGACACTCTTCAATGAGGCTCAACGACTTGTGTATTAATGTGAAATCCG 3660
 Db 3601 TGGAACTGTGACACTCTTCAATGAGGCTCAACGACTTGTGTATTAATGTGAAATCCG 3660
 Qy 3661 GAGGCCACTGTGGCCAGGAAATGTTTCAAGCATCCAAAGGCGCTGACATCTCT 3720
 Db 3661 GAGGCCACTGTGGCCAGGAAATGTTTCAAGCATCCAAAGGCGCTGACATCTCT 3720
 Qy 3721 GAGGACTCCCAAGGCTTTCGTCAAGTGTGAGAGTCAATGAGTGTGCTAGCTGTAC 3780
 Db 3721 GAGGACTCCCAAGGCTTTCGTCAAGTGTGAGAGTCAATGAGTGTGCTAGCTGTAC 3780
 Qy 3781 CAGGCGCAAGGCGGGAATGTGCTATGCTGAGCTCAAGAACATGCTGCTTCA 3840
 Db 3781 CAGGCGCAAGGCGGGAATGTGCTATGCTGAGCTCAAGAACATGCTGCTTCA 3840
 Qy 3841 CACTGCAAAAGCTCCCTGAGAGAAAGAACTGAAAGATGAACTGAACTCCAGCAT 3900
 Db 3841 CACTGCAAAAGCTCCCTGAGAGAAAGAACTGAAAGATGAACTGAACTCCAGCAT 3900
 Qy 3901 GGCATCTCCAGTTCCTCTACTGAGCAACATACACAGCGTGAAGACTTTCGCTTGTG 3960
 Db 3901 GGCATCTCCAGTTCCTCTACTGAGCAACATACACAGCGTGAAGACTTTCGCTTGTG 3960
 Qy 3961 GTGACGCTTTCCTTCCAAACACACTCACCTCCACTGAAAGAGAGGGAACACTGACTC 4020
 Db 3961 GTGACGCTTTCCTTCCAAACACACTCACCTCCACTGAAAGAGAGGGAACACTGACTC 4020
 Qy 4021 ACCTTCTTCGAGAGAGCTTTTCACTTCTGACACGCGGAGATGCGAGATGAGCATC 4080
 Db 4021 ACCTTCTTCGAGAGAGCTTTTCACTTCTGACACGCGGAGATGCGAGATGAGCATC 4080
 Qy 4081 GCACCTGGAACAACTGCTGGAACAGAGTGGCGGCAAGACTTCTCAACAACTTCAAC 4140
 Db 4081 GCACCTGGAACAACTGCTGGAACAGAGTGGCGGCAAGACTTCTCAACAACTTCAAC 4140
 Qy 4141 CACAGCGAGCCAACTCAAGTGCCTCTCTGAGAGCTTCACTTCAACCTTGC 4200
 Db 4141 CACAGCGAGCCAACTCAAGTGCCTCTCTGAGAGCTTCACTTCAACCTTGC 4200
 Qy 4201 AACAGCGAATGCTCCCAACAGAGTGAAGAACCCCGAGGTGCTTACTGAGGCTGTC 4260
 Db 4201 AACAGCGAATGCTCCCAACAGAGTGAAGAACCCCGAGGTGCTTACTGAGGCTGTC 4260
 Qy 4261 CCAGTGGCAGCGGAGTCCGCTTGTGTGAGGATCATCGGCAAGTGTGAGAGTGC 4320
 Db 4261 CCAGTGGCAGCGGAGTCCGCTTGTGTGAGGATCATCGGCAAGTGTGAGAGTGC 4320
 Qy 4321 AGGAGAGGTGGCGGAGGAAATCTTCAATGAGCTTGCACATGTGCGCTTAC 4377
 Db 4321 AGGAGAGGTGGCGGAGGAAATCTTCAATGAGCTTGCACATGTGCGCTTAC 4377

RESULT 2

US-10-094-749-390

; Sequence 390, Application US/10094749

; Publication No. US20030219741A1

; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: SOGIYAMA, TOMOYASU

; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO

```

; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKI, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKUJI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 390
; LENGTH: 2477
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-390

```

```

Query Match 22.2%; Score 970; DB 13; Length 2477;
Best Local Similarity 99.5%; Pred. No. 7.8e-259;
Matches 973; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1279 AGCGTCGGCGGAGATGGAACATGCGACCGTTACCAACCTTGGGAAACATCTCCGGAA 1338
DB 6 AGCGTCGGCGGAGATGGAACATGCGACCGTTACCAACCTTGGGAAACATCTCCGGAA 65
QY 1339 TTCACACCTTCCCTGAGAGGCTTCTGTTGGCACTGGGAAAGAACCAAGCTTAATGCC 1398
DB 66 TTCACACCTTCCCTGAGAGGCTTCTGTTGGCACTGGGAAAGAACCAAGCTTAATGCC 125
QY 1399 TTCTTAACCAAGCTGTGGAGAGGCGAGCTGAGATCTACCTGTCCAGGCCAGAGG 1458
DB 126 TTCTTAACCAAGCTGTGGAGAGGCGAGCTGAGATCTACCTGTCCAGGCCAGAGG 185
QY 1459 CTGTGTGACCTGATGAGATGAGACAGAGATACCTTCAGAGAACTGGAAATATA 1518
DB 186 CTGTGTGACCTGATGAGATGAGACAGAGATACCTTCAGAGAACTGGAAATATA 245
QY 1519 ACCCTGTTATAGGCGGCAATGACCTCTGTGATTTCTGCAATGATCTGTGCACTATTCT 1578
DB 246 ACCCTGTTATAGGCGGCAATGACCTCTGTGATTTCTGCAATGATCTGTGCACTATTCT 305
QY 1579 CCCGAGAACTTCAAGACAACTTGAAGAGCCCTGAGCACTCTCATGCTGAGGTTCT 1638
DB 306 CCCGAGAACTTCAAGACAACTTGAAGAGCCCTGAGCACTCTCATGCTGAGGTTCT 365
QY 1639 CGGGCAATTTGAACTGCTGTGAGCGGTGCTGTGATCTGCACTGAGAGAGCTTACCA 1698
DB 366 CGGGCAATTTGAACTGCTGTGAGCGGTGCTGTGATCTGCACTGAGAGAGCTTACCA 425
QY 1699 GAGAAAAAGTCTAATGAGGATGATCTGAGGTCTGAGTCTGCTGCTGCTGCTGAG 1758
DB 426 GAGAAAAAGTCTAATGAGGATGATCTGAGGTCTGAGTCTGCTGCTGCTGCTGAG 485
QY 1759 TTGTGATTAATCACTCAAGAACTTGTACCTCTCATGCAATTTCAACAAGATTTCAAGAG 1818
DB 486 TTGTGATTAATCACTCAAGAACTTGTACCTCTCATGCAATTTCAACAAGATTTCAAGAG 545
QY 1819 AAGACCAACCACTGATTTGAGAGTGGGCGATATGACAAAGGGAATTTTACTGTGTT 1878
DB 546 AAGACCAACCACTGATTTGAGAGTGGGCGATATGACAAAGGGAATTTTACTGTGTT 605
QY 1879 GTGAGCGGCTTCTTTGAAAACGTGGACATGCAAGAACCTCGAAGAGATTGCCGACAC 1938

```

```

DB 606 GTGAGCGGCTTCTTTGAAAACGTGGACATGCAAGAACCTCGAAGAGATTGCCGACAC 665
QY 1939 TCTTTCTGCTCTGACCTGTTTCCACTTCAGAGCAAGTCTTACCTCCGAGAGCCACT 1998
DB 666 TCTTTCTGCTCTGACCTGTTTCCACTTCAGAGCAAGTCTTACCTCCGAGAGCCACT 725
QY 1999 GCTCTCGGAACTAATGAGGAGGCTGTTGGCCAGAGAGCACTGTCATAAGTTGAA 2058
DB 726 GCTCTCGGAACTAATGAGGAGGCTGTTGGCCAGAGAGCACTGTCATAAGTTGAA 785
QY 2059 AACAGATCAATATCAATGTCGAAACAGAGTCCAGCCGTTTCTGAGACTTCAAGAAC 2118
DB 786 AACAGATCAATATCAATGTCGAAACAGAGTCCAGCCGTTTCTGAGACTTCAAGAAC 845
QY 2119 AGCATGAGGCTATGAGACCTGCTGCTGCAATGAGAGACAGAGCCCTTCTGCTTGCAC 2178
DB 846 AGCATGAGGCTATGAGACCTGCTGCTGCAATGAGAGACAGAGCCCTTCTGCTTGCAC 905
QY 2179 CCTACCTCAGTGCATGCGCTGAGACCTGAGACATCAAGTTGTGCTGCTGAGGAGAT 2238
DB 906 CCTACCTCAGTGCATGCGCTGAGACCTGAGACATCAAGTTGTGCTGCTGAGGAGAT 965
QY 2239 TCTGTGACCGCTGCAAT 2256
DB 966 TCTGTGACCGTGAAGACT 983

```

```

RESULT 3
US-09-764-891-1391
; Sequence 1391, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 1391
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (375)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (529)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-891-1391

```

```

Query Match 9.4%; Score 411.2; DB 11; Length 572;
Best Local Similarity 96.7%; Pred. No. 2.8e-120;
Matches 492; Conservative 2; Mismatches 6; Indels 9; Gaps 7;

```

```

QY 3042 CCTTTGAGCCAAATATGCTTGAACCACTTGGAGCAAAACAGAGACCTGTGAGAGC 3101
DB 15 CCTTTGAGCCAAATATGCTTGAACCACTTGGAGCAAAACAGAGACCTGTGAGAGC 74
QY 3102 AGAGATGCCCATCACTGTGCTCCACTAGAAATGAGCCCTTCTGAGAAACCTTGGAAATAG 3161
DB 75 AGAGATGCCCATCACTGTGCTCCACTAGAAATGAGCCCTTCTGAGAAACCTTGGAAATAG 134
QY 3162 TAACCTACGATACCCATCAAGCCAGAGCACTGAGAACTGGGCGAGTACTTCTGTGTAC 3221
DB 135 TAACCTACGATACCCATCAAGCCAGAGCACTGAGAACTGGGCGAGTACTTCTGTGTAC 194
QY 3222 AGAGTGAAGGCTTCAATAGTGTTCACCTCTGTGCAACAGCTCCGACAGACAGACAT 3281
DB 195 AGAGTGAAGGCTTCAATAGTGTTCACCTCTGTGCAACAGCTCCGACAGACAGACAT 254

```



```

APPLICANT: Kravkovskaya, L. L.
TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
FILE REFERENCE: 2760-103
CURRENT APPLICATION NUMBER: US/10/157,031
CURRENT FILING DATE: 2002-05-30
NUMBER OF SEQ ID NOS: 415
SOFTWARE: PatentIn version 3.1
SEQ ID NO 268
LENGTH: 1186
TYPE: DNA
ORGANISM: Homo sapiens
US-10-157-031-268

Query March          1.0%  Score 42;  DB 15;  Length 1186;
Beeb Local Similarity 50.5%;  Pred. No. 0.076;
Matches 102;  Conservative 0;  Mismatches 100;  Indels 0;  Gaps 0;

QY      2794  CTGACCTGTGGGGGAACTCCCAAGACTACCGAGGCTGTGAGGCTTTCAGCCGAGCTTAC 2853
Db      568  CCGCGGGTCTGGCGGGCTGCCCGCGGGGTCTACTGTGAGAGGGTCCACCGGCGAGCTG 627

QY      2854  CGAGCAGCATCGCGAGCTGTGGGGTCAGGCGGCTATGACAGCAGAGAGACTTCTCT 2913
Db      628  CTGCGCTGTCTGGCGGCACTTGCTGTGCCGACGCTGCTATCCCTTCAGAGCTGAGCCCG 687

QY      2914  GTGTGTCTGACAGCCCTTCTTCCAGAACATCCAGCTCCCTGTCTTCGGCGGATGGGCTCCA 2973
Db      688  GGCAGAGCTGTGGGCTGGTCTCCGCGGTGTGAGACCGCTGCTGCTGCAGAGGCTGGGCA 747

QY      2974  GATAGTCTCTCTTTGGCCCGAG 2995
Db      748  GACCAAGCCTTCAATTAGCCCTG 769

```

RESULT 6
US-10-157-031-268
; Sequence 268, Application US/10157031
; Publication No. US20030108690A1
; GENERAL INFORMATION:
; APPLICANT: BARNOVA, A. V.
; APPLICANT: YANKOVSKY, N. K.
; APPLICANT: KOZLOV, A. P.
; APPLICANT: LOBASHEV, A. V.

```

; APPLICANT: Hung, Mien-Chie
; APPLICANT: Lin, Shiao-Yih
; TITLE OF INVENTION: Methods and Compositions for Inhibiting EGF Receptor
; FILE REFERENCE: UTSC-720US
; CURRENT APPLICATION NUMBER: US/10/172,620
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/298,579
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 2643
; TYPE: DNA
; ORGANISM: Human
; US-10-172-620-15

```

```

Query Match      1.0%; Score 41.8; DB 15; Length 2643;
Best Local Similarity 61.5%; Pred. No. 0.15;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```

```

QY 4206 CCGATTGCTCCAGACAGGCTGAGAGCCCGGAGTGTCTACTGGGCTGTCCCACT 4265
DB 1993 CTGCGTACTTCCAGACAGGCTGTGTCTTCTCCATGACTCTGCGGAGGCTTTGA 1934
QY 4266 GGCAGCGGAGTGGCGCTTGTGTGGGCATCATCGGAGACAGTGTCTGG 4314
DB 1993 CGCAGTGGGGCCGTCAATGTACTGTGGGCACATGATACAGTTGTCTGG 1885

```

```

RESULT 9
US-09-725-433-1/c
; Sequence 1, Application US/09725433
; Patent No. US20020068362A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020068362A1artis AG
; TITLE OF INVENTION: Increased transgene expression in retroviral vectors having a sca
; FILE REFERENCE: 4-30921B/SYS
; CURRENT APPLICATION NUMBER: US/09/725,433
; CURRENT FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3633
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3633)
; OTHER INFORMATION:
; US-09-725-433-1

```

```

Query Match      1.0%; Score 41.8; DB 9; Length 3633;
Best Local Similarity 61.5%; Pred. No. 0.18;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```

```

QY 4206 CCGATTGCTCCAGACAGGCTGAGAGCCCGGAGTGTCTACTGGGCTGTCCCACT 4265
DB 1834 CTGCGTACTTCCAGACAGGCTGTGTCTTCTCCATGACTCTGCGGAGGCTTTGA 1775
QY 4266 GGCAGCGGAGTGGCGCTTGTGTGGGCATCATCGGAGACAGTGTCTGG 4314
DB 1774 CGCAGTGGGGCCGTCAATGTACTGTGGGCACATGATACAGTTGTCTGG 1726

```

```

RESULT 10
US-09-920-300A-1731/c
; Sequence 1731, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jianshun

```

```

; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121,547
; CURRENT APPLICATION NUMBER: US/09/920,300A
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 1789
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1731
; LENGTH: 5264
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2497
; OTHER INFORMATION: n = A,T,C or G
; US-09-920-300A-1731

```

```

Query Match      1.0%; Score 41.8; DB 10; Length 5264;
Best Local Similarity 61.5%; Pred. No. 0.22;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```

```

QY 4206 CCGATTGCTCCAGACAGGCTGAGAGCCCGGAGTGTCTACTGGGCTGTCCCACT 4265
DB 1753 CTGCGTACTTCCAGACAGGCTGTGTCTTCTCCATGACTCTGCGGAGGCTTTGA 1694
QY 4266 GGCAGCGGAGTGGCGCTTGTGTGGGCATCATCGGAGACAGTGTCTGG 4314
DB 1693 CGCAGTGGGGCCGTCAATGTACTGTGGGCACACTGATACAGTTGTCTGG 1645

```

```

RESULT 11
US-10-099-926-1731/c
; Sequence 1731, Application US/10099926
; Publication No. US2003016064A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jianshun
; APPLICANT: Secrist, Heather
; APPLICANT: Jianshun, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121,547C2
; CURRENT APPLICATION NUMBER: US/10/099,926
; CURRENT FILING DATE: 2002-03-17
; NUMBER OF SEQ ID NOS: 1982
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1731
; LENGTH: 5264
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2497
; OTHER INFORMATION: n = A,T,C or G
; US-10-099-926-1731

```

```

Query Match      1.0%; Score 41.8; DB 13; Length 5264;
Best Local Similarity 61.5%; Pred. No. 0.22;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```

```

QY 4206 CCGATTGCTCCAGACAGGCTGAGAGCCCGGAGTGTCTACTGGGCTGTCCCACT 4265
DB 1753 CTGCGTACTTCCAGACAGGCTGTGTCTTCTCCATGACTCTGCGGAGGCTTTGA 1694
QY 4266 GGCAGCGGAGTGGCGCTTGTGTGGGCATCATCGGAGACAGTGTCTGG 4314
DB 1693 CGCAGTGGGGCCGTCAATGTACTGTGGGCACACTGATACAGTTGTCTGG 1645

```

```

RESULT 12
US-10-033-528-1731/c

```

; Sequence 1731, Application US/10033528
; Publication No. US2002013197A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033.528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1731
; LENGTH: 5264
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2497
; OTHER INFORMATION: n = A,T,C or G
US-10-033-528-1731

Query Match 1.0%; Score 41.8; DB 14; Length 5264;
Best Local Similarity 61.5%; Pred. No. 0.22;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 4206 CCGATTGCTCCACAGACGAGCTGAAGAGCCCGAGAGTGTCTACTGGGCTGCCAGT 4265
Db 1753 CTGGTACTTCCAGACAGGAGGTGTGTTTCTCCATGACTCCGCGGACAGGTCTTGA 1694

Qy 4266 GCGAGCGGAGTGGCGCTTGTGTGGGCATCATCGGAGGAGTGTCTGG 4314
Db 1693 CGCAGTGGGCGCGCTCATGTAGTGGGACACTGATACAGTTGTCTGG 1645

RESULT 13
US-10-007-926A-137/c

; Sequence 137, Application US/10007926A
; Publication No. US2003014353A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOUTGATE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007.926A
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 137
; LENGTH: 5532
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: epidermal growth factor receptor (avian
; OTHER INFORMATION: erythroid-leukemia viral (v-erb-b) oncogene
; OTHER INFORMATION: homology) (EGFR) gene.
US-10-007-926A-137

Query Match 1.0%; Score 41.8; DB 13; Length 5532;
Best Local Similarity 61.5%; Pred. No. 0.23;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 4206 CCGATTGCTCCACAGACGAGCTGAAGAGCCCGAGAGTGTCTACTGGGCTGCCAGT 4265
Db 2020 CTGGTACTTCCAGACAGGAGGTGTGTTTCTCCATGACTCCGCGGACAGGTCTTGA 1961

Db 2020 CTGGTACTTCCAGACAGGAGGTGTGTTTCTCCATGACTCCGCGGACAGGTCTTGA 1961
Qy 4266 GCGAGCGGAGTGGCGCTTGTGTGGGCATCATCGGAGGAGTGTCTGG 4314
Db 1960 CGCAGTGGGCGCGCTCATGTAGTGGGACACTGATACAGTTGTCTGG 1912

RESULT 14
US-10-101-510-95/c

; Sequence 95, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: MAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101.510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 95
; LENGTH: 5532
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-95

Query Match 1.0%; Score 41.8; DB 13; Length 5532;
Best Local Similarity 61.5%; Pred. No. 0.23;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 4206 CCGATTGCTCCACAGACGAGCTGAAGAGCCCGAGAGTGTCTACTGGGCTGCCAGT 4265
Db 2020 CTGGTACTTCCAGACAGGAGGTGTGTTTCTCCATGACTCCGCGGACAGGTCTTGA 1961

Qy 4266 GCGAGCGGAGTGGCGCTTGTGTGGGCATCATCGGAGGAGTGTCTGG 4314
Db 1960 CGCAGTGGGCGCGCTCATGTAGTGGGACACTGATACAGTTGTCTGG 1912

RESULT 15
US-10-380-931-17/c

; Sequence 17, Application US/10380931
; Publication No. US20030215944A1
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: C. Frank Bennett
; APPLICANT: Jacqueline Wyatt
; APPLICANT: Susan M. Freier
; TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HER-1 EXPRESSION
; FILE REFERENCE: RTSP-0187
; CURRENT APPLICATION NUMBER: US/10/380.931
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: 09/676,610
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 182
; SEQ ID NO 17
; LENGTH: 5532
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (187)...(3819)
US-10-380-931-17

Query Match 1.0%; Score 41.8; DB 13; Length 5532;
Best Local Similarity 61.5%; Pred. No. 0.23;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 4206 CCGATTGCTCCACAGACGAGCTGAAGAGCCCGAGAGTGTCTACTGGGCTGCCAGT 4265
Db 2020 CTGGTACTTCCAGACAGGAGGTGTGTTTCTCCATGACTCCGCGGACAGGTCTTGA 1961

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 06:59:00 ; Search time 236 Seconds
(without alignments)
8186.161 Million cell updates/sec

Title: US-10-054-691-1

Perfect score: 4377
Sequence: 1 atcgggcctgcggccagcat.....tgcacatgtgcccctctag 4377

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: Issued Patents NA:*
2: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
3: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
5: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
6: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	41.8	1.0	5532	2 US-08-475-035-3	Sequence 3, Appl1
C 2	41.8	1.0	5532	4 US-09-676-6108-17	Sequence 17, Appl1
C 3	41.8	1.0	169998	4 US-09-676-6108-24	Sequence 24, Appl1
C 4	41.8	1.0	197486	4 US-09-877-1774-10	Sequence 10, Appl1
5	40.8	0.9	4689	3 US-09-105-537-34	Sequence 34, Appl1
6	40.8	0.9	36778	3 US-09-105-537-5	Sequence 5, Appl1
7	40.8	0.9	38506	3 US-09-320-878-19	Sequence 19, Appl1
8	40.8	0.9	38506	4 US-09-141-908-1	Sequence 1, Appl1
9	40.8	0.9	38506	4 US-09-657-440-19	Sequence 19, Appl1
10	40.2	0.9	1465	4 US-09-338-671-1	Sequence 1, Appl1
11	39.6	0.9	800	5 PCT-US95-04801-4	Sequence 1, Appl1
12	38	0.9	7218	1 US-08-232-463-14	Sequence 14, Appl1
13	37.6	0.9	513	4 US-09-252-991A-5791	Sequence 5791, Ap
14	37.6	0.9	762	4 US-09-252-991A-5809	Sequence 5809, Ap
15	37.6	0.9	1077	4 US-09-252-991A-5839	Sequence 5839, Ap
16	37.6	0.9	2007	4 US-09-252-991A-5770	Sequence 5770, Ap
17	37.2	0.8	840	4 US-09-252-991A-1634	Sequence 1634, Ap
18	37.2	0.8	1335	4 US-09-252-991A-973	Sequence 973, Ap
19	37.2	0.8	1335	4 US-09-252-991A-1685	Sequence 1685, Ap
20	37.2	0.8	1335	4 US-09-252-991A-1111	Sequence 1111, Ap
21	37.2	0.8	4563	4 US-09-252-991A-930	Sequence 930, App
22	37.2	0.8	1413	4 US-09-252-991A-6071	Sequence 6071, Ap
23	37	0.8	1431	4 US-09-252-991A-6232	Sequence 6232, Ap
24	37	0.8	1506	4 US-09-252-991A-6228	Sequence 6228, Ap
25	37	0.8	2005	4 US-09-636-872A-18	Sequence 18, Appl1
26	37	0.8	2005	4 US-09-636-872A-18	Sequence 18, Appl1
27	36.8	0.8	1555	2 US-08-696-376-1	Sequence 1, Appl1

28	36.8	0.8	3150	3 US-08-943-768-1	Sequence 1, Appl1
29	36.8	0.8	3150	4 US-09-865-960-1	Sequence 1, Appl1
30	36.6	0.8	2502	1 US-08-073-384C-7	Sequence 7, Appl1
31	36.6	0.8	2502	1 US-08-254-359A-7	Sequence 7, Appl1
32	36.6	0.8	2502	1 US-08-483-043-7	Sequence 7, Appl1
33	36.6	0.8	2502	1 US-08-481-238-7	Sequence 7, Appl1
34	36.6	0.8	2502	2 US-08-471-066B-7	Sequence 7, Appl1
35	36.6	0.8	2502	2 US-08-484-956-7	Sequence 7, Appl1
36	36.6	0.8	2502	2 US-08-559-491-7	Sequence 7, Appl1
37	36.6	0.8	2502	2 US-08-756-386-7	Sequence 7, Appl1
38	36.6	0.8	2502	2 US-08-823-516-7	Sequence 7, Appl1
39	36.6	0.8	2502	3 US-08-682-853A-7	Sequence 7, Appl1
40	36.6	0.8	2502	3 US-08-758-314-7	Sequence 7, Appl1
41	36.6	0.8	2502	3 US-08-758-314-7	Sequence 7, Appl1
42	36.6	0.8	2502	4 US-09-350-309-7	Sequence 7, Appl1
43	36.6	0.8	2502	4 US-08-520-946-7	Sequence 7, Appl1
44	36.6	0.8	2502	4 US-09-684-938-7	Sequence 7, Appl1
45	36.6	0.8	2502	4 US-09-684-938-7	Sequence 7, Appl1

ALIGNMENTS

RESULT 1
US-08-475-035-3/C
Sequence 3, Application US/08475035

Patent No. 5985553

GENERAL INFORMATION:

APPLICANT: KING, C. R.

APPLICANT: KRAUS, MATTHIAS H.

APPLICANT: AARONSON, STUART A.

TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM

TITLE OF INVENTION: BGF RECEPTOR GENE

NUMBER OF SEQUENCES: 4

CORRESPONDING ADDRESSES:

ADDRESSER: NEEDLE & ROSENBERG, P.C.

STREET: Suite 1200, 127 Peachtree Street

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30303

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/475,035

FILING DATE: 7 Jun 1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Perryman, David G.

REGISTRATION NUMBER: 33,438

REFERENCE/DOCKET NUMBER: 1414.656

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404/688-0770

TELEFAX: 404/688-9880

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 5532 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 187..3816

US-08-475-035-3

Query Match

Best Local Similarity 61.5%;

Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Score 41.8; DB 2; Length 5532;

Pred. No. 0.27;

0; Mismatches 42; Indels 0; Gaps 0;

```

RESULT 2
US-09-676-610B-17/c
: Sequence 17, Application US/09676610B
: Patent No. 6444465
: GENERAL INFORMATION:
: APPLICANT: C. Frank Bennett
: APPLICANT: Jacqueline Wyatt
: APPLICANT: Susan M. Freier
: TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HER-1 EXPRESSION
: FILE REFERENCE: RTS-0138
: CURRENT APPLICATION NUMBER: US/09/676,610B
: CURRENT FILING DATE: 2000-09-29
: NUMBER OF SEQ ID NOS: 182
: SEQ ID NO 17
: LENGTH: 5532
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (187)...(3819)
: US-09-676-610B-17

```

```

? RESULT 3
? US-09-676-6108-24/C
? Sequence 24, Application US/09676610B
? Patent No. 6444465
? GENERAL INFORMATION:
? APPLICANT: C. Frank Bennett
? APPLICANT: Jacqueline Wyatt
? APPLICANT: Susan M. Freier
? TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HER-1 EXPRESSION
? FILE REFERENCE: RTS-0138
? CURRENT APPLICATION NUMBER: US/09/676,610B
? CURRENT FILING DATE: 2000-09-29
? NUMBER OF SEQ ID NOS: 182
? SEQ ID NO 24
? LENGTH: 165998
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: exon
? LOCATION: (1208)...(1472)
? NAME/KEY: intron
? LOCATION: (1473)...(124390)
? NAME/KEY: exon
? LOCATION: (124391)...(124544)
? NAME/KEY: intron
? LOCATION: (124545)...(125409)
? NAME/KEY: exon
? LOCATION: (125410)...(125595)
? NAME/KEY: intron
?

```

LOCATION: (125596) ... (128711)	NAME/KEY: exon
LOCATION: (128712) ... (128848)	NAME/KEY: int3on
LOCATION: (128849) ... (133400)	NAME/KEY: exon
LOCATION: (133401) ... (133469)	NAME/KEY: int3on
LOCATION: (133470) ... (134652)	NAME/KEY: exon
LOCATION: (134653) ... (134773)	NAME/KEY: int3on
LOCATION: (134774) ... (136116)	NAME/KEY: exon
LOCATION: (136117) ... (136361)	NAME/KEY: int3on
LOCATION: (136362) ... (137936)	NAME/KEY: exon
LOCATION: (137937) ... (138053)	NAME/KEY: int3on
LOCATION: (138054) ... (138637)	NAME/KEY: exon
LOCATION: (138638) ... (138766)	NAME/KEY: int3on
LOCATION: (138767) ... (138864)	NAME/KEY: exon
LOCATION: (138865) ... (138940)	NAME/KEY: int3on
LOCATION: (138941) ... (139765)	NAME/KEY: exon
LOCATION: (139766) ... (139860)	NAME/KEY: int3on
LOCATION: (139861) ... (142245)	NAME/KEY: exon
LOCATION: (142246) ... (142445)	NAME/KEY: int3on
LOCATION: (142446) ... (143605)	NAME/KEY: exon
LOCATION: (143606) ... (143738)	NAME/KEY: int3on
LOCATION: (143739) ... (145838)	NAME/KEY: exon
LOCATION: (145839) ... (145931)	NAME/KEY: int3on
LOCATION: (145932) ... (147385)	NAME/KEY: exon
LOCATION: (147386) ... (147544)	NAME/KEY: int3on
LOCATION: (147545) ... (153274)	NAME/KEY: exon
LOCATION: (153275) ... (153321)	NAME/KEY: int3on
LOCATION: (153322) ... (155088)	NAME/KEY: exon
LOCATION: (155089) ... (155231)	NAME/KEY: int3on
LOCATION: (155232) ... (156025)	NAME/KEY: exon
LOCATION: (156026) ... (156151)	NAME/KEY: int3on
LOCATION: (156152) ... (156826)	NAME/KEY: exon
LOCATION: (156827) ... (156928)	NAME/KEY: int3on
LOCATION: (156929) ... (163399)	NAME/KEY: exon
LOCATION: (163400) ... (163386)	NAME/KEY: int3on

Query Match	1.0%;	Score 41.8;	DB 4;	Length 169998;
Best Local Similarity	61.5%;	Pred. No. 2.1;		
Matches 67;	Conservative 0;	Mismatches 42;	Indels 0;	Gaps 0;

Query Match	1.0%;	Score 41.8;	DB 4;	Length 169998;
Best Local Similarity	61.5%;	Pred. No. 2.1;		
Matches 67;	Conservative 0;	Mismatches 42;	Indels 0;	Gaps 0;

Query Match	1.0%;	Score 41.8;	DB 4;	Length 169998;
Best Local Similarity	61.5%;	Pred. No. 2.1;		
Matches 67;	Conservative 0;	Mismatches 42;	Indels 0;	Gaps 0;

QY 4206 CCATTGCTCCAGACCAAGGCTTGAAGAGCCCGAGTCTCTACTGAGCTGTCCAGT 4265
Db 147498 CTGCGTACTTCCAGACCAAGGCTGTGTTTCTCCAGTACTCTGCGGAGAGTCTTGA 147439
QY 4266 GCGAGGCGGAGTGGCTTGTGTGTCATCATCGGAGACAGTGTCTGG 4314
Db 147438 CGCAGTGGGGGCGTCAATGATGAGGACACATGATACAGTGTCTGG 147390

RESULT 4

US-09-877-177A-10/c
; Sequence 10, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; TITLE OF INVENTION: And Correlation of Levels thereof With Survival
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 197496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-10

Query Match 1.0%; Score 41.8; DB 4; Length 197496;
Best Local Similarity 61.5%; Pred. No. 2.3;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 4206 CCATTGCTCCAGACCAAGGCTTGAAGAGCCCGAGTCTCTACTGAGCTGTCCAGT 4265
Db 155498 CTGCGTACTTCCAGACCAAGGCTGTGTTTCTCCAGTACTCTGCGGAGAGTCTTGA 155439
QY 4266 GCGAGGCGGAGTGGCTTGTGTGTCATCATCGGAGAGTGTCTGG 4314
Db 155438 CGCAGTGGGGGCGTCAATGATGAGGACACATGATACAGTGTCTGG 155390

RESULT 5

US-09-105-537-34
; Sequence 34, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438051
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 4689
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-34

Query Match 0.9%; Score 40.8; DB 3; Length 4689;
Best Local Similarity 45.8%; Pred. No. 0.47;
Matches 141; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 3986 TCACCCCACTGAAGAGAGGAGACCTGACCTTCTTCTCGAGAGATGTTTC 4045
Db 2588 TCACCTTACTCGCGAGAGCTTGAGCCACGCTTCACTGAGAGATGAGCCATCTCC 2647
QY 4046 ACTTCTCAGACCGCGGAGATGCGAGATGCGCATTCGACTCTGGAACATATCTGGAC 4105

Db 2648 CCAACCGCAACCGGCAACACCCCGAGCTCCCACTTACGCTTCCAGACGAGCTTCT 2707
QY 4106 CAGTGGCGGCAAGACTACTCTCAACAACTTACACCCAGAGCGGAGCAACTGAAGTCC 4165
Db 2708 GAGTGAAGAGCTCCGCGCCACACAGCGCGGAGAGCACTGGGCTTACCGGCTGATGGA 2767
QY 4166 CCTTCTCAGAGACCTTACTCTTACACCTTGGGAAACAGCGATTCTTCCAGACGAG 4225
Db 2768 AGCGCTGACGGGCTCCGAGCAGGCGGAGCTTGTCCGGGCGTGTATGTCGCGTGGGA 2827
QY 4226 CTGAAGAGACCCCGAGTGTCTTACTGAGCTGTCCCAATGAGAGCGGAGTGGCTTG 4285
Db 2828 GCGAGCCAGAGCGGAGCTGTGGGCGCGCTGAAGCGCGGAGCGAGTCACTGATAC 2887
QY 4286 TGGTGGGC 4293
Db 2888 TGGAGCC 2895

RESULT 6

US-09-105-537-5
; Sequence 5, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438051
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 36778
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-5

Query Match 0.9%; Score 40.8; DB 3; Length 36778;
Best Local Similarity 45.8%; Pred. No. 1.6;
Matches 141; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 3986 TCACCCCACTGAAGAGAGGAGACACTGACTTCTTCTTCTCCGAGAGTGTTC 4045
Db 29578 TCACCTTACTCGCGAGAGCTTGAGCCACGCTTCACTGAGAGTGGGCGCCATCTCC 29637
QY 4046 ACTTCTCAGACCGCGGAGATGCGAGATGCGCATTCGAGACATATCTGGAGC 4105
Db 29638 CACCGCAACCGGCAACACCCGAGCTCCCACTTACGCTTCCAGACGAGCGCTTCT 29697
QY 4106 CAGTGGCGGCAAGACTACTCTCAACAACTTACACCCAGAGCGGAGCAACTGAAGTCC 4165
Db 29698 GAGTGAAGAGCTCCGCGCCACACAGCGCGGAGAGACTGGGTTACCGGCTGATGGA 29757
QY 4166 CCTTCTCAGAGACCTTACTCTTACACCTTGGGAAACAGCGATTCTTCCAGACGAG 4225
Db 29758 AGCGCTGACGGGCTCCGAGCAGGCGGAGCTTGTCCGGGCGTGTATGTCGCGTGGGA 29817
QY 4226 CTGAAGAGACCCCGAGTGTCTTACTGAGCTGTCCCAATGAGAGCGGAGTGGCTTG 4285
Db 29818 GCGAGCCAGAGCGGAGCTGTGGGCGCGCTGAAGCGCGGAGCGAGTCACTGATAC 29877
QY 4286 TGGTGGGC 4293
Db 29878 TGGAGCC 29885

RESULT 7

US-09-320-878-19
; Sequence 19, Application US/09320878A
; Patent No. 6117659

```

; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; CURRENT FILING DATE: 1999-05-27
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908
; EARLIER FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: 60/119,139
; EARLIER FILING DATE: 1999-02-08
; EARLIER APPLICATION NUMBER: 60/100,880
; EARLIER FILING DATE: 1998-09-22
; EARLIER APPLICATION NUMBER: 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-320-878-19

```

```

Query Match          0.9%; Score 40.8; DB 3; Length 38506;
Best Local Similarity 45.8%; Pred. No. 1.7;
Matches 141; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

```

```

QY 3886 TCACCCCACTGAACGAGAGAGGAGACACTGACTGACTCTTCTTCCGAGACTGTTTC 4045
    |||||
DB 27720 TCACCTCACTGCGCGAAGCCCTGGGCGCAACGCGCTTACATCGATGGCGCCCATCTCC 27779
    |||||
QY 4046 ACTTCTCAGACCGCGGCGATGCGAGATGCGCATGCGACTGGAACAACATGCTGGAAC 4105
    |||||
DB 27780 CCACCGCAACCGGCGACACCCCGAGCTCCCACTTCCAGACGAGCGCTTCT 27839
    |||||
QY 4106 CAGTGGCGCGCAAGACTTCACTTCAACAATTACCCACGCGGAGCCAACTCAAGTCC 4165
    |||||
DB 27840 GCGTGCAGAGACTCGGCGCCACGACGCGCGCGGAGCGGCTTACCGCTCAAGTGA 27899
    |||||
QY 4166 CCTTCTCTGAGAGCCCTTACTCTTACACCTGCGGAAACGCGATTGCTCCAGACCAAG 4225
    |||||
DB 27900 AGCCGCTGACGCGCTCCGCGGCGAGCGGACCTGTCCGCGCGGTGATCGTCCGCGGA 27959
    |||||
QY 4226 CTGAAGAAGCCCGGAGGTGCTTACTGAGGCTGTCCAGTGGCAGCGGAGTCCGCTTG 4285
    |||||
DB 27960 GCGAGCCAGAAAGCGAGCTGTGGGCGCGCTGAAGCGCGGAGCGGAGTGAAGTAC 28019
    |||||
QY 4286 TGGTGGGC 4293
    |||||
DB 28020 TGGAAAGCC 28027
    |||||

```

```

RESULT 8
US-09-141-908-1
; Sequence 1, Application US/09141908
; Patent No. 6503741
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
; FILE REFERENCE: 300622002100
; CURRENT APPLICATION NUMBER: US/09/141,908

```

```

; CURRENT FILING DATE: 1998-08-28
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538
; EARLIER FILING DATE: 1998-05-06
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247
; EARLIER FILING DATE: 1997-04-30
; EARLIER APPLICATION NUMBER: PROV. 60/076,919
; EARLIER FILING DATE: 1998-03-05
; EARLIER APPLICATION NUMBER: PROV. 60/087,080
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-141-908-1

```

```

Query Match          0.9%; Score 40.8; DB 4; Length 38506;
Best Local Similarity 45.8%; Pred. No. 1.7;
Matches 141; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

```

```

QY 3886 TCACCCCACTGAACGAGAGAGGAGACACTGACTGACTCTTCTTCCGAGACTGTTTC 4045
    |||||
DB 27720 TCACCTCACTGCGCGAAGCCCTGGGCGCAACGCGCTTACATCGATGGCGCCCATCTCC 27779
    |||||
QY 4046 ACTTCTCAGACCGCGGCGATGCGAGATGCGCATGCGACTGGAACAACATGCTGGAAC 4105
    |||||
DB 27780 CCACCGCAACCGGCGACACCCCGAGCTCCCACTTCCAGACGAGCGCTTCT 27839
    |||||
QY 4106 CAGTGGCGCGCAAGACTTCACTTCAACAATTACCCACGCGGAGCCAACTCAAGTCC 4165
    |||||
DB 27840 GCGTGCAGAGACTCGGCGCCACGACGCGCGCGGAGCGGCTTACCGCTCAAGTGA 27899
    |||||
QY 4166 CCTTCTCTGAGAGCCCTTACTCTTACACCTGCGGAAACGCGATTGCTCCAGACCAAG 4225
    |||||
DB 27900 AGCCGCTGACGCGCTCCGCGGCGAGCGGACCTGTCCGCGCGGTGATCGTCCGCGGA 27959
    |||||
QY 4226 CTGAAGAAGCCCGGAGGTGCTTACTGAGGCTGTCCAGTGGCAGCGGAGTCCGCTTG 4285
    |||||
DB 27960 GCGAGCCAGAAAGCGAGCTGTGGGCGCGCTGAAGCGCGGAGCGGAGTGAAGTAC 28019
    |||||
QY 4286 TGGTGGGC 4293
    |||||
DB 28020 TGGAAAGCC 28027
    |||||

```

```

RESULT 9
US-09-657-440-19
; Sequence 19, Application US/09657440
; Patent No. 6509455
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/657,440
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141,908
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-657-440-19

```

```

Query Match          0.9%; Score 40.8; DB 4; Length 38506;

```

Best Local Similarity 45.8%; Pred. No. 1.7;
Matches 141; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 3986 TCACCCACTGAGACGAGAGAGGGGACACTGACCTCCTCTTCCGAGAGACTTTTC 4045
Db 27720 TCACCTCAGCTGCGGAAAGCTTGCGCAACGCGCTCAGCTCAGATGGGCGCCCTCTCC 27779
QY 4046 ACTTCTCAGACCGCGGCGATGCCAGATGCGCATCGCACTCTGGAACAATGCTGGAA 4105
Db 27780 CCACCGGAACGGCGCACCAACCCGAGCTCCCACTACGCTTCCAGACCGAGCTTCT 27839
QY 4106 CAGTGGCGCGCAAGACTACTCTCAACAATTCAACCAACCGGAGCCAACTCAAGTCC 4165
Db 27840 GCGTGCAGAGCTCGCGCCCAACGCGCCGCGAGCACTGAGCTTACCGGCTGAGTGA 27899
QY 4166 CCTCTCTGAGAGCCCTTACTCTTACACCTCTGCGGAAACAGCCGATTTCTCCAGACGAG 4225
Db 27900 AGCCGCTGAGCGGCTCTCGGCGCAGCGGAGCTCTGCGGCGGATGCTGCGCTCGGGA 27959
QY 4226 CTGAAGAGCGCGCGAGGCTCTTACTGAGCTGTCCAGTGGCGAGGAGTGGCGCTTG 4285
Db 27960 GCGAGCGAGAGCGCGAGCTCTGAGCGCGCTGAGAGCGCGGAGGAGGAGTGCAGCTAC 28019
QY 4286 TGGTGGGC 4293
Db 28020 TGGAGGCC 28027

RESULT 10

US-09-338-671-1
; Sequence 1, Application US/09338671
; Patent No. 6194638
; GENERAL INFORMATION:
; APPLICANT: Dhugga, Kanwarpal
; APPLICANT: Fallis, Patricia Lynne
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Alteration of Hemicellulose
; FILE REFERENCE: 0782
; CURRENT APPLICATION NUMBER: US/09/338, 671
; CURRENT FILING DATE: 1999-06-22
; EARLIER APPLICATION NUMBER: 60/090,416
; EARLIER FILING DATE: 1998-06-23
; NUMBER OF SEQ. ID NOS.: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ. ID NO. 1
; LENGTH: 1451
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (100)...(1194)
US-09-338-671-1

Query Match 0.9%; Score 40.2; DB 3; Length 1451;
Best Local Similarity 54.4%; Pred. No. 0.34; Indels 0; Gaps 0;

Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 2905 GACTTCTGTGTGTGTGAGAGCCCTTCTTCAGAACTCCAGTCCCTGTCTGGCGGAT 2964
Db 196 GACTTCTGTGAGATGTGGCGGCTTCTTCAGAGCTTACCACTCATCATGTGTGAGAC 255
QY 2965 GGGCTCCCAATATGCTCTTCTTGGCCGAGCTGATCCCAATATCAAAATTTGAC 3024
Db 256 GGGGACCGGACCAAGACCATCAAGTGTCCGAGGGCTTGTGATGAACTTCAACCGC 315
QY 3025 TCCAGCTGGCGAGAGCCCTTTGGACCA 3053
Db 316 AAGACATCAACCGCATCTCTGGGCGCA 344

RESULT 11
PCT-US95-04801-4

; Sequence 4, Application PC/TUS9504801
; GENERAL INFORMATION:
; APPLICANT: Martin, Juan F.
; APPLICANT: Coque, Juan R.
; APPLICANT: Enguita, Francisco J.
; APPLICANT: Fuente, Juan L.
; APPLICANT: Lirio, Francisco J.
; APPLICANT: Lirio, Paloma
; TITLE OF INVENTION: DNA ENCODING CEPHARYCIN BIOSYNTHESIS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07065

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04801
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wallen III, John W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-4720
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ. ID NO.: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 800 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US95-04801-4

Query Match 0.9%; Score 39.6; DB 5; Length 800;
Best Local Similarity 50.8%; Pred. No. 0.36; Indels 3; Gaps 1;
Matches 121; Conservative 0; Mismatches 114; Indels 3; Gaps 1;

QY 2691 CAACCTGCTGAGCTTCTTCTGAAACCCCACTATCATGCGGAGTGTCTCTGGAAACCCAGA 2750
Db 95 CACCGCGAGAGACTTCTGTGACCTC---AACCTGTTCCGGGGGCTGGGGAGAGACCCGCT 151
QY 2751 CAAGTCCGAGTGCAGAGCGGCGAGCGCTTTGTGTAACTGTGACCTTGGCGGAGAA 2810
Db 152 CTACCAACCCCGCGTGTGCGAGCGCGCGCGGCGGAGCTGCGCTGACCGTGGCGCA 211
QY 2811 CTCCCAAGAGTACGCGAGGCTTGAAGGCTTCAAGCCGAGCTTACCGAGAGCATGCCGA 2870
Db 212 GGCCTCCGCGGATCTGGGTTCTCCGACTTGGCGCTTACAGATGGGCGGCTGCGCAT 271
QY 2871 GCTGTGGGAGTGCAGCGGCTATGACAGCGAGAGACTTCTGTGTGTGCTGAGAGCC 2928
Db 272 GCTGAAGAACCGGAGACCGGCGGCTTACAGACCTATGATGTGAACTTGGCGGCC 329

RESULT 12
US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 INMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base paire
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpc-Fls
US-08-232-463-14

Query Match 0.9%; Score 38; DB 1; Length 7218;
Best Local Similarity 4.5%; Pred. No. 3.8;
Matches 17; Conservative 197; Mismatches 162; Indels 0; Gaps 0;
QY 3479 GGGAGGGGACGAGCACTAATGTGGACGCGGAGGCGCAGCTAGGACATGCCAG 3558
DB 1410 RR 1351
QY 3539 CCCAGGCTGGACCTGTAGAGCAATGAAAGAGCCCGACATCAACCTGGAGAAAG 3598
DB 1350 RR 1291
QY 3599 ACTGGAAGCTGGTCACTCTTCAATTGGGGTCAACGACTGTGTCAATTAGTGAATC 3658
DB 1290 RR 1231
QY 3659 CGGAGCGCCACTTGGCAGCAGATATGTCAGCATCCACAGCCCTGGACATCTCT 3718
DB 1230 RR 1171
QY 3719 CTGAGAGCTCCCAAGGCTTTCGTCACGTGTGAGAGTCATGAGCTGAGCTGACCTGT 3778
DB 1170 RR 1111
QY 3779 ACCGAGCGCAAGCGGAAATGTCCATGCTGCGACCTAGCAACAATGCACTTGCTCA 3838
DB 1110 RR 1051
QY 3838 GACACTGGCAAGCTC 3854
DB 1050 ACTGACCGCCAAAGCTC 1035

RESULT 13
US-09-252-991A-5791
Sequence 5791, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 5791
LENGTH: 513
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5791

Query Match 0.9%; Score 37.6; DB 4; Length 513;
Best Local Similarity 49.5%; Pred. No. 1;
Matches 97; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 2790 CGTTTGACCTGGGGAGAACTCCCAAGAGCTAGCCAGGCTGAGGCTTGACCGAGC 2849
DB 101 CGCCCTGTGGCTGGATCATCGCCCTGGCGGCTGCTGTGCTGCGGCTTACGCCAC 160
QY 2850 CTACCGGAGAGCATGCGCGAGCTGTGGGATCAGCCGCTATGACAGGAGAGACTT 2909
DB 161 CGTCCGAAACGCTGTATCGGGATCGCACCGCTAGTGAATCTTCGCAACGT 220
QY 2910 CTGTGTGTGTGAGGCTTCTTCCAGAAATCAGCTCCCTGCTCGGCGATGGGCT 2969
DB 221 GCCGCTGTGTGAGCTGTTCATCTGTTCTCTGTACTCCGACCTGTGCCGAGGG 280
QY 2970 CCCAGATAGCTCTTC 2985
DB 281 CTGACAGACTGTTC 296

RESULT 14
US-09-252-991A-5809
Sequence 5809, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 5809
LENGTH: 762
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5809

Query Match 0.9%; Score 37.6; DB 4; Length 762;
Best Local Similarity 49.5%; Pred. No. 1.3;
Matches 97; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 2790 CGTTTGACCTGGGGAGAACTCCCAAGAGCTAGCCAGGCTTGACCGAGC 2849
DB 124 CGCCCTGTGGCTGGATCATCGCCCTGGCGGCTGCTGTGCTGCGGCTTACGCCAC 183

QY 2850 TTACCGGAGCAGTATGCGCAGCTGTGGGGGTACGCGCTATGACACGACGAGAGACTT 2909
 Db 184 CTTGCCGAAACCGCCTGTATGCGGGATGCCACGCGCTACGTGGAACCTTCCGCAAGT 243
 QY 2910 CTCTGTGTGCTGCAGCCCTCTTCACAGACATCCAGCTCCCTGTCCTGGCGGATGGCT 2965
 Db 244 GCCCGCTGTGTGACGCTGTTCATCTGTACTTCTGTATACCGACCTGTCGCCGAGGG 303
 QY 2970 CCCGATACGTCCTTC 2985
 Db 304 CTTGCAACCTGTTC 319

RESULT 15

```

US-09-252-991A-5839/C
; Sequence 5839, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5839
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-5839

```

Query Match	0.98	Score 37.6	DB 4	Length 1077
Best Local Similarity	49.58	Pred.No. 1.6		
Matches 97	Conservative	0	Mismatches 99	Indels 0
			Gaps	0

QY	2790	CGTTCTAACCCTGGGGGAGAACTCCCAAGAGTACCAAGGCTGGAGGCTTCAGCGAGC	2849
Db	649	CGCCCTGGTCGGCTGGATATCGCCCTTGCGCGTGGGCTTCGCTGCTGGCGGTATGCGAC	590
QY	2850	CTACCGGAGCAGATGCGCGAGCTGGTGGGGTCAAGCGCCTATGACACGACGAGAGACTT	2909
Db	589	CGTCCGGAACGGCTGGTATCGGGGATGCGCACCGCTACGTGGAATCTTCCGAACT	530
QY	2910	CTCTGTGTGCTGCAGGCCCTTCTTCAGAAATCCAGCTCCCTGTCTCTGGCGGATGGCT	2968
Db	529	GCCCTGTGTGGAGCTGTTCATCTGTACTTCTGTACCCGACCTGTGCCGAGG	470
QY	2970	CCCGATACGTCCTTC	2985
Db	469	CTGCGAGACTGTGTC	454

Search completed: January 6, 2004, 14:45:59
Job time : 243 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 06:31:30 ; Search time 8433 Seconds
(without alignments)
12614.810 Million cell updates/sec

Title: US-10-054-691-1
Perfect score: 4377
Sequence: 1 atcggggctgcggcagcat.....tgcgcactgtgcacctctag 4377

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estp1:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gse_hum:*
18: em_gse_inv:*
19: em_gse_pln:*
20: em_gse_vrt:*
21: em_gse_fun:*
22: em_gse_mam:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rod:*
26: em_gse_dhg:*
27: em_gse_vrt1:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1626	37.1	3148	11	AK030142 Mus muscu
2	868.8	19.8	935	9	AL542856 AL542856
3	788	18.0	1081	12	BM923066 BM923066
4	766.4	17.5	935	13	BU150503 BU150503

C 5	760.6	17.4	903	13	EX349938
C 6	732.2	16.7	974	13	EX349937
C 7	730.6	16.7	900	13	EX369782
C 8	712.6	16.3	861	13	BU183160
C 9	708.2	16.2	880	13	BU1912651
C 10	687.8	15.7	1186	13	EX354650
C 11	677.4	15.5	1033	12	BM549919
C 12	671.4	15.3	763	12	BI912073
C 13	627.8	14.3	823	12	BI488437
C 14	607	13.9	607	12	BM714799
C 15	578	13.2	1562	11	AK015993
C 16	576.2	13.2	608	10	AW851123
C 17	467.8	10.7	587	9	AW752833
C 18	450.2	10.3	617	9	AI424952
C 19	419.8	9.6	978	13	BU540700
C 20	405	9.3	716	13	BU683311
C 21	401	9.2	625	14	CA447656
C 22	398.8	9.1	412	12	BM146939
C 23	372.2	8.5	583	9	AA18082
C 24	371.6	8.5	389	12	BM146940
C 25	365.4	8.3	616	9	AW467395
C 26	349.6	8.0	979	14	BU715569
C 27	344.4	7.9	570	10	BF510463
C 28	343	7.8	567	13	BU739708
C 29	339.4	7.8	575	13	BU743739
C 30	322.2	7.4	648	14	BY731604
C 31	322.2	7.4	2929	11	AK087596
C 32	317.4	7.3	541	10	BF854199
C 33	310.2	7.1	701	12	BI696117
C 34	303.8	6.9	667	10	BB616890
C 35	301.8	6.9	525	10	BE137831
C 36	301.4	6.9	525	10	BE684057
C 37	299.8	6.8	469	14	CB733804
C 38	299.8	6.8	513	4	BX517602
C 39	299.6	6.8	511	10	BE855594
C 40	294.4	6.7	511	10	BE645470
C 41	284.4	6.5	430	9	AA466625
C 42	282.2	6.4	495	9	AA461827
C 43	274.6	6.3	409	9	AA512437
C 44	263.4	6.0	725	13	BU633808
C 45	262.8	6.0	762	10	BF168176

ALIGNMENTS

RESULT 1
AK030142
LOCUS
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4932703L18 product:weakly similar to PHOSPHOLIPASE B [Rattus norvegicus], full insert sequence.
ACCESSION AK030142
VERSION AK030142.1 GI:26081683
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carinci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
2 Carinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL 2049374
MEDLINE 11042159
PUBMED

- REFERENCE
AUTHORS
3
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kikunishi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, M.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujieko, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsui, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-364-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
- JOURNAL
MEDLINE
20530913
11076861
- TITLE
4
Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Akakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Konno, S., Yamakawa, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, D., Mombaerts, P., Nordone, P.,
Ringo, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Sessa, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.,
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
- JOURNAL
MEDLINE
21085660
11217851
- TITLE
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
- JOURNAL
REFERENCE
AUTHORS
6 (bases 1 to 3148)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Haneigaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numata, R., Ono, M., Ohnato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
- FEATURES

source	
1. .3148	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="FANTOM_DB:4932703L18"
	/db_xref="taxon:10090"
	/clone="4932703L18"
	/sex="male"
	/issue_type="testis"
	/clone_id="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="adult"
misc_feature	
1. .3148	
/note="weakly similar to PHOSPHOLIPASE B [Rattus norvegicus] (SPTRL054728, evidence: FASTY, 84%id, 26.6%length, match=1158)"	
BASE COUNT	
856 a 819 c 756 g 717 t	
ORIGIN	
Query Match 37.1%; Score 1626; DB 11; Length 3148;	
Best local Similarity 75.7%; Pred. No. 0;	
Matches 2074; Conservative 0; Mismatches 640; Indels 24; Gaps 4;	
52	CAAGGACCCCTCAGATCATACCTCTCTAGAAAGATGATGGAAGGCGACTATGG 111
82	CAAGGGCCCTCCAAATCCATGAGCTCTTCTGAGAGAACATATGGCATGCCAG 141
112	CCAGA-----GACCTGAAAGATTCCTCCATGCCATGCAACCAATTAATAGAGTG 165
142	CAAGTGTTTGGACCCCTGAAGAAATTTCCATTCCTTGCAAGCCAAAGAGTTAGAAATG 201
166	AATATGCTTCTAATACAGTTCAGTCTCTGAGACCTTGTATTAATTTGGAGAGCC 225
202	AGGTGCTTTCTAGTCAAGTCACTCTCTGAGACCTCTGAGACCTTAATCTGGCAGCC 261
226	ATTGGCAATGTGAATTCCTCCAGACCCAGGAGCGGCGATGTGGAAG----- 276
262	ATCGGCAACCAAAATTCCTTAGCCCTGGGCTGGCAGATCAACATGAGAAACT 321
277	CAAGACTGAGCTGAAGGCCAGCAGAGTGTGATGGAGTGTGACAGCTTTTACAG 336
322	CAAGACATCAAGAACCAACCAAGAGATGATGATGGAGATCAATGACATTTTACAG 361
337	ATCATCAGATATTTACGCTCTCTGTTCCATGCTCTGTGCGCACTGGAAGAGAGTC 396
382	ATCATCAGATATTTACGCTCTCTGTTCCATGCTCTGTGCGCACTGGAAGAGAGTC 441
397	ATACCCAGAGTGTGCTGAAGCTTGTGATTCAGGCTCAAGACTGTGAGAAACATG 456
442	GCAGTCCACACTACTGCTGAAGATTTGTGGATTCAAGCTTAAGAGCTGGAGCGCCTG 501
457	AAAGAGAACTGCACTTGAATTTGATTTGATGAGAGTCAATGATGTTCTTCACT 516
502	AAGGACAAACCCGAACTGACTTTGAGAGAGCTGGAATTCATCACTGTTCTTCACT 561
517	AATGCAAGCAGATGTTACCTGCTCCCTGCTCAACAGAAATGGGCTTGGGCGGCGC 576
562	AACACAGACCAAGTCACTGCTGTCTCTCTCTGCTCAGCAGAAAGTCACTGATGAGGAC 621
577	GTCAGATGAGTGTGAGGAGTGTGCTGAGATCACTGCTCAGCAGAGAGTCCCAAGAGATTGTA 636
622	ATGAGATGCTGTGAGGAGTGTGAGATTCATCATCATGATGAGTCCCAAGAGATTGTTG 681
637	AACCTGTCAGCTCTCTGAGGTTGAGAGGTTCTCTGTCATGATCAACGACACTTGGCTC 636
682	AATTTGATGATCTCTCTGAGGTTTGGCCATGATCTTCAGATCAAGAGACTTGGCTTC 741
697	AGCCCTGACACAGAGCCCTGATTTGCTCAGAGAGACCAACCGGCTGGCAGAGTGGTG 756
742	AGCCCTGACACAGAGGTTTGAATGACAC--GAGACACAGACGCTATCAAGAGCTGTC 798
757	ATCAGATGCTTATAGAGAGCTGGAACAGCTCTCTGCTCCAGCAGATGACAGTGA 816
799	ATCAGATGCTTATAGAGAGCTGGAAGATCTTCTGCTCTGACCAAGTTCAACAG 858

QY 817 CAGAGTCCCTTCAACCGTGTGTTTCAAGCCTTTCTGATGAGACCAACCCATCTTACAC 876
 DB 859 CATTGACCTTGGCGGTGGTTTTCAGACCTTTCTTGATGATAATTGAAAC-----ACCC 912
 QY 877 TGGAGAGACCCCGACTCCAGATTTTCAACGCTGGCTGGGACTCTTGGAAATGAGATG 936
 DB 913 TTGAAAAGGTCCCTGGCCCCAGGATCCCAACACTGGCCTTCAGGATCTTGGATGACATG 972
 QY 937 ATGGAGCCAGCAGGAGAGAAAGATGAGCCATTGAGTGTAAACAACGGAGGCCAATGAAAG 996
 DB 973 ATGGAACCAAGTGGGTCAAAAAGATGGGCTACTCAATATGACAGAGAGAAAGAAATGAAA 1032
 QY 997 TGTCCCTCTCAGAGAGACCCCTATCTGTTCAGCTTACAGAAACAGCAACTAGCAACGA 1056
 DB 1033 TGTCCCTCTGAGGAGAGTCCCTATCTGTTCACCTTCAAAAATGCAACTACAGGCCAGA 1092
 QY 1057 CTGCAGAAACCCCAAGACAACTTGAAGTGAAGAAAGAGCGGAAATCAGATGCTGAC 1116
 DB 1093 CGACTGAAACCTATTAACAAAGCTTCAGATGAAAGAAAGATCAGAAATTCACCTGCTGAC 1152
 QY 1117 AAAGACCCCTCCGATAGGTTCCACCTCAGTTCATAGGTGAAGCCGGCTGACATCAAC 1176
 DB 1153 AAGAACCCCTCGAACTCAATCCCAACAGCTTCAGAGCTTGGCGCAGCTGATATCAAG 1212
 QY 1177 GTAATTGAGACCCCTGGGTGACTCTCTCAGCGCAGGCAATGGGGCCGGGTCAACCTGGG 1236
 DB 1213 ATCATAGAGACATGGGTGATCTGCTCAGCGCAGGCAATGAGCTGGGCGACCCCTTGG 1272
 QY 1237 AAGCTTGGAGCGTCTTGACTCAGTACGAGGCGCTGCTGAGCGTGGCGGAGATGAG 1296
 DB 1273 AATATCTTGGATGTCTTAATGAAATCCAGGCGCTGCTGAGTGGCGGAGATGAG 1332
 QY 1297 AACATGGGACCGTTACCACTCGGGAACATCTCTCGGGAATTCACCTTCCCTGAAG 1356
 DB 1333 ACCATGAAAAACCGTGAACACCTCAACCAATCTCCGGGAATTCACACCTCTTGAAG 1392
 QY 1357 GGGCTCTCTGTTGGCACTGGGAAAGAAACAGTCTTAATGCTCTTAAACAGGCGTGG 1416
 DB 1393 GGGCTTCTGTTGGCACTGGGAAAGAAAGCACTTCCGAGACATCTTCAACAGGCGGTA 1452
 QY 1417 GCAGAGGCGCAGCTGAGATCTACCTGTTCAGAGGCCAGGCGTGGTGAACCTGATGAG 1476
 DB 1453 GCAGAGGCCAAATCTGATGGCTTAACTGCTGGCAGGCCAGAAAGTGGTGAACCTGATGAG 1512
 QY 1477 AATGACACAGAGATCACTTTCAGAAAGATGGAAGATTAATCCTGTTTATAGCGGC 1536
 DB 1513 GCAGACAAAGCAATTAACCTTTCAGAAAGATGGAAGATTAATCACTGTTTATAGAGGC 1572
 QY 1537 AATGACCTCTGATTTCTGCAATGATCTGCTCAATTTCTCCAGAACTTCAAGAC 1596
 DB 1573 AATGACCTCTGCTTCTGCAATTAATCTGATCTGCTTCTCCCAAACTTCAATAGAC 1632
 QY 1597 AACATTGAAAGGCCCTGGAATCTCTCAATGCTGAGAGTTCTCGGGCACTTGTGAACCTG 1656
 DB 1633 AACATGAAGAGGCCCTGGAATCTCTCAATGAGAGTTCTCGGGCACTTGTGAACCTG 1692
 QY 1657 GTGACGCTGCTGAGATCTGCAACTGAGGAGCTGTACAGAGAGAAAAAGTCTACTGC 1716
 DB 1693 GCATGTGTATGATGATCAACCCCTTGAGAGAACTATTCAATGAACCTACAGTCTATGC 1752
 QY 1717 CCAAGATATGCTCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1776
 DB 1753 CCAAGATATGCTCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1812
 QY 1777 GAACCTGCTACCTCTGATTCGAATTCGAAGAGTTTCAAGAGAGAACCCCACTGAT 1836
 DB 1813 GAACCTTCCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1872
 QY 1837 GAGAGTGGCGATATGACACAGGAGAAATTTTACTGTGTGTTGACCGCTTCTTGA 1896
 DB 1873 AATAGCGGACGATATGACACAGGAGAAATTTTACTGTGTGTTGACCGCTTATTTGA 1932

QY 1897 AACGTGACATGCAAGACCTTCGAGAGATTTGCTGACATCTTCTGCTGAC 1956
 DB 1933 AATGTGTCATGCAACGACCCCGAGAGGCGTGGCCGACAAAGTCTTCTTGGCCGAC 1992
 QY 1957 TGTTCCTTACAGACAGATCTCACTCCGAGCAGCGATGCTCTTGGAAACAATATG 2016
 DB 1993 TGTTCCTTACATTAATGCAAGATCTCACTGCGCTGAGCCATGCTCTGGAAGAACATG 2052
 QY 2017 CTGAGCCCTGTTGGCCAGAGACGACTGCTCAATGATTGAAAACAAGATCAATATAC 2076
 DB 2053 CTGGAACCTGTGGCCCAAGACAGACATTAATTTTGAATCAAGGCCCTATGCTG 2112
 QY 2077 TGTCCGACAGGCTCCAGCGCTTCTGAGACCTTACAGAACAGCATGACGGTCAATGG 2136
 DB 2113 TGTCCGACAGGCTCCAGCGCTTCTGAGACCTTACAGAACAGCATGACGGTCAATGG 2172
 QY 2137 ACTGGCTGCTGACAGGACAGAGCCCTTCTGCTTGAACCTTACCTGATGACATGCC 2196
 DB 2173 ACTGGATGTGTGCTGAGAGAGAGCCCTTCTGCTTCAACCAACCTGATGACATACC 2232
 QY 2197 CTGAGACCTGACAGACATCCAGATGAGTGGCTGCTGAGGAGATTTCTGACCGCTGCAAT 2256
 DB 2233 CTGAGACCTGACAGACATTCAGATGAGTGGCTGCTGAGAGACTTCTGACCTGCTCAAT 2292
 QY 2257 GGAATTTGGCTCCAAACCAAGACCTCCCGATGTCACCAACAGATGCGGAGCTGCA 2316
 DB 2293 GGAATTTGGCTCCAAAGATTAATCTCACCGATGTCACCAACAGATGCGGAGCTGCA 2352
 QY 2317 TACAGTGAAGAGAGGAGCGCTCCCTGAGAAATGAGACCACTTACCTTAATCTTGG 2376
 DB 2353 TATAGTGTGTGGAGAGAGAGCCCTGAGAAATGAGACCACTTGGCCATATCTTACGG 2412
 QY 2377 GAGTTTAAAGAAACCTCACAGGCTTACCGCTGAGGAGAGGATGACAGAT 2436
 DB 2413 AATTTTAAAGAAATCTCACAGGCTTACCGATGAGAAACCGGATCTCAGTTCCGAAAC 2472
 QY 2437 GCAATCTCAATCAAGCTGTTCCCGAGCAAGAGCTGAGATTTTGAAGCCAAATGCA 2496
 DB 2473 GCTTCTCTTAACAGGCTGCTCCTGGGCGGAGGCTGAGAACCTTCAAGCCAAATGCGG 2532
 QY 2497 ACTGTGATGACAGAGATGAAGATGATCATAGATTAATTTCCATGAAGACTGGAAGTTC 2556
 DB 2533 ACTGTGATGACAGAGATGAAGATGATCATAGATTAATTTCCATGAAGACTGGAAGTTC 2592
 QY 2557 ATCAAGTGTGTGATGAGAGGAGCCGATTTATGATGATCTGACAGATTCGATCTGAT 2616
 DB 2593 ATCAAGTGTGTGATGAGAGGAGCCGATTTATGATGATCTGACAGATTCGATCTGAT 2652
 QY 2617 TGTGACGCAACTTTGTTGATGACCATCTGCGCAATGCTTGGAGCTGCTGATGAGAGTTC 2676
 DB 2653 TGTGACGCAACTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2712
 QY 2677 CCCAGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2736
 DB 2713 CCCAGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2772
 QY 2737 CTGGAAGAACCCAGACAAAGTGGCCAGTGAATCAGTTCAG 2774
 DB 2773 CTGGAAGAACCCAGACAAAGTGGCCAGTGAATCAGTTCAG 2810
 RESULT 2
 AL542856
 LOCUS 935 bp mRNA linear EST 12-MAY-2003
 DEFINITION AL542856 Homo sapiens PLACBNTA Homo sapiens cDNA clone CS080E013YE22
 5-PRIME, mRNA sequence.
 ACCESSION AL542856
 VERSION AL542856.2 GI:30548429
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 935)
 AUTHORS Li, W. B., Gruber, C., Jeessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT On Feb 15, 2001 this sequence version replaced gi:12875311.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 1474.f
 Contact: Feng Liang Email: fliang@life.techno.com URL:
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID: CS0DE013BC110P1.
 FEATURES
 source
 1..935
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE013YE23"
 /tissue_type="PLACENTA"
 /clone_id="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo (dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."
 BASE COUNT 229 a 272 c 262 g 172 t
 ORIGIN
 Query Match 19.8%; Score 868.8; DB 9; Length 935;
 Best Local Similarity 98.1%; Pred. No. 7.5e-215;
 Matches 879; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
 Oy 3477 CTGGAGAGGAGACAGACGACTAAATGTGGACGCGAAGGGCCAGAGCTAGAGCATGCC 3536
 Db 40 CTGGTACCGGTCCGGAATTCCTCGGATGCGACGAGGGCCAGAGCTAGAGCATGCC 99
 Oy 3537 AGCCAGAGGCTGGGACCTGTGTAGAGCGAATGAAAAGCCCGACATCACTCGGAGAA 3596
 Db 100 AGCCAGAGGCTGGGACCTGTGTAGAGCGAATGAAAAGCCCGACATCACTCGGAGAA 159
 Oy 3597 AGACTGAGAGCTGTGACACTCTTCATTTGGGGTCAACGACTTGTGATTAAGTGAAGA 3656
 Db 160 AGACTGAGAGCTGTGACACTCTTCATTTGGGGTCAACGACTTGTGATTAAGTGAAGA 219
 Oy 3657 TCCGAGAGGCCACTTGGCCAGCAAGATATGTTACAGACATCCACAGGCGCTGACATCCT 3716
 Db 220 TCCGAGAGGCCACTTGGCCAGCAAGATATGTTACAGACATCCACAGGCGCTGACATCCT 279
 Oy 3717 CTCTGAGAGAGCTCCCAAGGGGTTTCGTCAAGTGTGTGAGGTCTAGAGCTGGTACGCT 3776
 Db 280 CTCTGAGAGAGCTCCCAAGGGGTTTCGTCAAGTGTGTGAGGTCTAGAGCTGGTACGCT 339
 Oy 3777 GTACACAGGCGCAAGGCGGAAATGTGCATCTGGCAGCTCAGAACTGCACTTGCTCT 3836
 Db 340 GTACACAGGCGCAAGGCGGAAATGTGCATCTGGCAGCTCAGAACTGCACTTGCTCT 399
 Oy 3837 CAGACACTCGCAAGGCTCCCTGAGAGCAAGAACTGAAGAACTGAAGCTTGAACCTTCCA 3896
 Db 400 CAGACACTCGCAAGGCTCCCTGAGAGCAAGAACTGAAGAACTGAAGCTTGAACCTTCCA 459
 Oy 3897 GCATGCGATCTCCAGTTTCTCTTACTGGACCAATACACAGAGGTGAGGACTTTGGGCT 3956
 Db 460 GCATGCGATCTCCAGTTTCTCTTACTGGACCAATACACAGAGGTGAGGACTTTGGGCT 519
 Oy 3957 TGTGCTGAGAGCTTCTTCCAAAACACACTCACCCTGAAGCAGAGAGGAGGACACTGA 4016
 Db 520 TGTGCTGAGAGCTTCTTCCAAAACACACTCACCCTGAAGCAGAGAGGAGGACACTGA 579
 Oy 4017 CTTGACCTTCTTCTCGAGAGACTGTTTCACTTCTCAGACCGCGGCGCATGCCAGATGCC 4076
 Db 580 CTTGACCTTCTTCTCGAGAGACTGTTTCACTTCTCAGACCGCGGCGCATGCCAGATGCC 639

Oy 4077 CATGCACTCTGGAACACATGCTGGAACAGTGGGCGCGAAGACTACTCAACAACCTT 4136
 Db 640 CATGCACTCTGGAACACATGCTGGAACAGTGGGCGCGAAGACTACTCAACAACCTT 699
 Oy 4137 CACCACAGCCGAGCCCAACTCAAGTCCCTCTCTGAGAGGCTTACTCTTACACCTT 4196
 Db 700 CACCACAGCCGAGCCCAACTCAAGTCCCTCTCTGAGAGGCTTACTCTTACACACCT 759
 Oy 4197 GCGGAACAGCCGATTTGCTCCCAACAGGCTGAAGAGCCCGAGAGTGTCTTACTGCGG 4256
 Db 760 GCGGAACAGCCGATTTGCTCCCAACAGGCTGAAGAGCCCGAGAGTGTCTTACTGCGG 819
 Oy 4257 TGTCCAGTGGCAGCGGAGTCCGCTTGTGTGGGCAATCATCGGACAGTGTCTGAG 4316
 Db 820 TGTCCAGTGGCAGCGGAGTCCGCTTGTGTGGGCAATCATCGGACAGTGTCTGAG 879
 Oy 4317 GTCCAGAGAGGTGGCCGAGGAGAGATCTCCATGAGGCTTCCGACCTGTGAGCC 4372
 Db 880 GTCCAGAGAGGTGGCCGAGGAGAGATCTCCATGAGGCTTCCGACCTGTGAGCC 935
 RESULT 3
 BM923066
 LOCUS
 DEFINITION AGENCOURT_6632017 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5757033
 5' mRNA sequence.
 ACCESSION BM923066
 VERSION BM923066.1 GI:19373445
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1081)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-rc@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.liml.gov
 Plate: LHM42798 row: j column: 10
 High quality sequence stop: 664.
 Location/Qualifiers
 1..1081
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5757033"
 /tissue_type="leukocyte"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_118"
 /note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH MGC Library."
 BASE COUNT 255 a 329 c 286 g 208 t 3 others
 ORIGIN
 Query Match 18.0%; Score 788; DB 12; Length 1081;
 Best Local Similarity 93.1%; Pred. No. 9.7e-194;
 Matches 925; Conservative 0; Mismatches 52; Indels 17; Gaps 9;

```

QY 2648 ATGCTTGGACGTCCTGATAGAGAGGTGGCCAGAGTCTGTGCAACCTGTGAGCTTCC 2707
Db 16 ATGCTTGGACGTCCTGATAGAGAGGTGGCCAGAGTCTGTGCAACCTGTGAGCTTCC 75
QY 2708 TGAACCCCACTATCATGCGGAGGTGTCTCTGGGAAACCAAGACAGTCCCACTGAGC 2767
Db 76 TGAACCCCACTATCATGCGGAGGTGTCTCTGGGAAACCAAGACAGTCCCACTGAGC 135
QY 2768 AGGCGACGCTTGTGTATGCTGCTGACCCCTGGGAGAACTCCCAAGAGCTTACCA 2827
Db 136 AGGCGACGCTTGTGTATGCTGCTGACCCCTGGGAGAACTCCCAAGAGCTTACCA 195
QY 2828 GGGCTGAGGCTTTCAGCCGAGCTTACCGAGACAGTCCGAGAGCTGTGTGAGGCTC 2887
Db 196 GGGCTGAGGCTTTCAGCCGAGCTTACCGAGACAGTCCGAGAGCTGTGTGAGGCTC 255
QY 2888 GCTATGACACGACGAGAGGACTTCTGTGTGCTGACGCCCTTCTTCAGAACTCCAGC 2947
Db 256 GCTATGACACGACGAGAGGACTTCTGTGTGCTGACGCCCTTCTTCAGAACTCCAGC 315
QY 2948 TCCCTGCTCTGGGGAGAGTGGCTCCCAATAGTCTTCTTGGCCAGACTGATCCACC 3007
Db 316 TCCCTGCTCTGGGGAGAGTGGCTCCCAATAGTCTTCTTGGCCAGACTGATCCACC 375
QY 3008 CAATCAGAAATTCACCTCCAGCTGCGACAGAGCCCTTTGAGCAATATGCTTGAACAC 3067
Db 376 CAATCAGAAATTCACCTCCAGCTGCGACAGAGCCCTTTGAGCAATATGCTTGAACAC 435
QY 3068 TTGGAAGCAAAACAGAGACCTCTGAGACCTGAGAGAGATGCCATGCTGTCCACTC 3127
Db 436 TTGGAAGCAAAACAGAGACCTCTGAGACCTGAGAGAGATGCCATGCTGTCCACTC 495
QY 3128 AGAATGAGCCCTTCTCTGAGAAACCTCTGGAAATAGTAACTACAGTACCCATCAAGCAG 3187
Db 496 AGAATGAGCCCTTCTCTGAGAAACCTCTGGAAATAGTAACTACAGTACCCATCAAGCAG 555
QY 3188 CCAATTGAGACTGGGAGTGAATCTCTGTGTACAGAGTGAAGGCTTCCAAATGATGTTG 3247
Db 556 CCAATTGAGACTGGGAGTGAATCTCTGTGTACAGAGTGAAGGCTTCCAAATGATGTTG 615
QY 3248 CAACCTCTGTCCACACAGCTCCGACAGACAGATCAATCAAGTGTGGCCG-CCCTGGGTGAC 3306
Db 616 CAACCTCTGTCCACACAGCTCCGACAGACAGATCAATCAAGTGTGGCCG-CCCTGGGTGAC 675
QY 3307 TCTCTGACTACAGAGTGGGAGTCTGACCAACCACTCTCACTGATCCATCTTGG 3366
Db 676 TCTCTGACTACAGAGTGGGAGTCTGACCAACCACTCTCACTGATCCATCTTGG 735
QY 3367 AGGGAAGCTCTCTTGGAGCACTTGGAGGGGATGGGAATTTGAGACTCACACC--ACACTGC 3424
Db 736 AGGGAAGCTCTCTTGGAGCACTTGGAGGGGATGGGAATTTGAGACTCACACCACACTGC 795
QY 3425 CCAACATTTCTGAAGAAGTTCAACCTT--ACCTCTTGGCTTCTCTACAGACCTGGG- 3481
Db 796 CCAACATTTCTGAAGAAGTTCAACCTTCTTACTCTTCTTGGGCTTCTCTACAGACCTGGG 855
QY 3482 AGGAGGACAGAG-ACCTAAATGTGG--CAAGCGAAGGGGCGAAGTGAAGG-ACATGCC 3536
Db 856 AGGAGGACAGAG-ACCTAAATGTGGCAACGGGAAGGGGCGAAGTGAAGG-ACATGCC 915
QY 3537 AGCCCGAGCC--TGGACCTGTGTAGAGCGAATGAAAA--ACAGCCCCACATCAACT 3590
Db 916 AGCCCGAGCCCTGTGGGAGCTGTGGTAAGCCGATGAAAAACAGCCCCGAAATCAACCT 975
QY 3591 GGAGAAAGACTGGAAGCTGTGTCACTTTCATT 3624
Db 976 GGAGAAAGACTGTAAACTGGGTCAAACTTCTTT 1009

```

RESULT 4
BU150503 935 bp mRNA linear EST 03-SEP-2002

```

DEFINITION AGENCOURT_8125945_Lupeki_dorsal_root_ganglion_Homo_sapiens_cDNA
ACCESSION BU150503
VERSION BU150503.1 GI:22664035
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC. http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strauberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM1355 row: n column: 23
High quality sequence start: 53
High quality sequence stop: 532.
Location/Qualifiers
1. 935
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:617622"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_1b="Lupski_dorsal_root_ganglion"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
Note: Site 2: Sali; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACCGCTCCG-3' and
5'-GACTAGTCTTAGATCGGAGCGGCGCCCTT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
BASE COUNT 235 a 290 c 248 g 162 t
ORIGIN
Query Match 17.5%; Score 766.4; DB 13; Length 935;
Best Local Similarity 95.0%; P-Val: 3.7e-188;
Matches 857; Conservative 0; Mismatches 36; Indels 9; Gaps 6;
QY 3484 GGGACAGCAGAGTAAATGTGCGAGCGGAGGGCCAGAGCTAGGACATGCCAGCCAG 3543
Db 1 GGGACAGCAGAGTAAATGTGCGAGCGGAGGGCCAGAGCTAGGACATGCCAGCCAG 60
QY 3544 GCTTGGAGCTGTGAGAGCGAATGAAAAACAGCTCCGACATCAACTGAGAAAGACTGG 3603
Db 61 GCTTGGAGCTGTGAGAGCGAATGAAAAACAGCTCCGACATCAACTGAGAAAGACTGG 120
QY 3604 AACCTGTCACTCTTCACTTGGGGTCAAGACTGTGTATTAATCTGTGAATCCGGAG 3663
Db 121 AACCTGTCACTCTTCACTTGGGGTCAAGACTGTGTATTAATCTGTGAATCCGGAG 180
QY 3664 GCCCACTTGGCCAGGAATATGTTACAGACATCCAAAGGCCCTTGACATCTCTGAG 3723
Db 181 GCCCACTTGGCCAGGAATATGTTACAGACATCCAAAGGCCCTTGACATCTCTGAG 240
QY 3724 GAGCTCCCAAGGGCTTTCGTCAGCTGTGTGAGAGTCAATGAGTGTGCTGATCCAG 3783
Db 241 GAGCTCCCAAGGGCTTTCGTCAGCTGTGTGAGAGTCAATGAGTGTGCTGATCCAG 300

```

QY 3784 GGCCAAAGCGGGAATGTCATGCTGACAGCTCAGAACACTGCTGCTCAGACAC 3843
 Db 301 GGCCAAAGCGGGAATGTCATGCTGACAGCTCAGAACACTGCTGCTCAGACAC 360
 QY 3844 TCGCAAGCTCCTCGAGAGAGCAAGAACTGAGAAAGTGAACCTTCACATGCG 3903
 Db 361 TCGCAAGCTCCTCGAGAGAGCAAGAACTGAGAAAGTGAACCTTCACATGCG 420
 QY 3904 ATCTCCAGTTTCTCTCTGCTGACCAATPACACAGAGGTAGAGCTTGGGGTGTG 3963
 Db 421 ATCTCCAGTTTCTCTCTGCTGACCAATPACACAGAGGTAGAGCTTGGGGTGTG 480
 QY 3964 CAGCCTTCTTCCAAACACACTACCCCACTGAACAG-AGAGGGGACACTGACTAC 4022
 Db 481 CAGCCTTCTTCCAAACACACTACCCCACTGAACAGAGGGGACACTGACTAC 540
 QY 4023 CTCTTCTCGAGAGCTGTTTTCATCTTCTGACACCGGGGCTGAGATGGCATGCG 4082
 Db 541 CTCTTCTCGAGAGCTGTTTTCATCTTCTGACACCGGGGCTGAGATGGCATGCG 600
 QY 4083 ACTCTGAAACAATGCTGGAACCAAGTGGGCGCGGAACCTTCCCAACCTTCAACCA 4142
 Db 601 ACTCTGAAACAATGCTGGAACCAAGTGGGCGCGGAACCTTCCCAACCTTCAACCA 660
 QY 4143 CAGCGGAGCCAACTCAAGTGCCTCTCTGAGAGCCCTTACTACACCTGCGG-A 4201
 Db 661 CAGCGGAGCCAACTCAAGTGCCTCTCTGAGAGCCCTTACTACACCTGCGGAA 720
 QY 4202 ACAAGCATTGCTCCAGACCAAGCTGAGAGAGCCCGAGAGTCTTACTGAGGCTGTC 4261
 Db 721 ACAAGCATTGCTCCAGACCAAGCTGAGAGAGCCCGAGAGTCTTACTGAGGCTGTC 780
 QY 4262 CAGTGGGA--GCGGAGTCCGCTTGTGTGGGATCATCGGCACT-AGTCTGAGAGT 4318
 Db 781 CAGGAGGAGAGGAGAACTCCGCTTGTGTGGGATCATCTTGTGAAAGTGGGCTGAGAGT 840
 QY 4319 GCAGAGAGAGTGG---CCGAGAGAGAGATCTCTCA--TGAGCCCTGCGCACTGTGGCCTC 4374
 Db 841 GCAGAGAGAGTGGCCCGCCCTGTGTGGGATCATCTTGTGAAAGTGGGCTGAGAGT 900
 QY 4375 TA 4376
 Db 901 CA 902
 RESULT 5
 BX349938/c 903 bp mRNA linear EST 05-MAY-2003
 LOCUS BX349938 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 DEFINITION CDNA clone CS0DC023YA08 3-PRIME, mRNA sequence.
 ACCESSION BX349938
 VERSION BX349938.1 GI:30371420
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 903)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 1474.f,
 Contact: Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0BA10052A12_CS00473_2.
 Location/Qualifiers
 1..903
 FEATURES
 source /organism="Homo sapiens"

/mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC023YA08"
 /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-0.190 (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 162 a 242 c 269 g 223 t 7 others
 ORIGIN
 Query Match 17.4%; Score 760.6; DB 13; Length 903;
 Best Local Similarity 92.6%; Pred. No. 1.2e-186;
 Matches 833; Conservative 0; Mismatches 26; Indels 41; Gaps 2;
 QY 3431 TTCTGAAGAAATTCAACCTTACCTCTTGGCTTCTCTACACAGCTGGAGGGGACAG 3490
 Db 903 TTGGAAGAAATTCAACCTTACCTCTTGGCTTCTCTACACAGCTGGAGGGGACAG 844
 QY 3491 CAGGACTAAATGTGGCAGAGGGGAGGGGAGGAGCTAGGGGACATGCCAGCCCTGGG 3550
 Db 843 CAGGACTAAATGTGGCAGAGGGGAGGGGAGGAGCTAGGGGACATGCCAGCCCTGGG 784
 QY 3551 ACCTGGTAGAGCAATGAAAAACAGCCCGACATCAACTGAGAAAGACTGAGAGCTGG 3610
 Db 783 CCTGGTAGAGCAATGAAAAACAGCCCGACATCAACTGAGAAAGACTGAGAGCTGG 724
 QY 3611 TCACACTCTTCAATTGGGGTCAACGACTGTGTCTATTACTGTGAAATCC----- 3659
 Db 723 TCACACTCTTCAATTGGGGTCAACGACTGTGTCTATTACTGTGAAATCCGGGAGCCCC 664
 QY 3660 -----GAGAGCCCACTGGCCAGCGAATATGTTTCAG 3690
 Db 663 GACCAACCCCATGGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
 QY 3691 CACATCCAAAGAGCCCTGAGCATCTCTCTGAGAGAGCTCCCAAGGCTTCTGCAACGTG 3750
 Db 603 CACATCCAAAGAGCCCTGAGCATCTCTCTGAGAGAGCTCCCAAGGCTTCTGCAACGTG 544
 QY 3751 GTGGAGGTCAATGAGCTGGCTAGCTGTACCAAGGSCCAAGCGGGAATGTGCAATGCTG 3810
 Db 543 GTGGAGGTCAATGAGCTGGCTAGCTGTACCAAGGSCCAAGCGGGAATGTGCAATGCTG 484
 QY 3811 GCAGCTCAGAAACAATGCACTTGGCTTCAAGCACTCCCAAAAGCTCCCTGAGAGCAAGAA 3870
 Db 483 GCAGCTCAGAAACAATGCACTTGGCTTCAAGCACTCCCAAAAGCTCCCTGAGAGCAAGAA 424
 QY 3871 CTGAAGAAAGTGAAGTGAAGCTCCAGCATGAGCATCTCCAGTTTCTCTACTGSCACAA 3930
 Db 423 CTGAAGAAAGTGAAGTGAAGCTCCAGCATGAGCATCTCCAGTTTCTCTACTGSCACAA 364
 QY 3931 TACACAAGCGTGAAGACTTTGGGTTGTGTGACAGCTTTTCTTCCAAACACACTACCC 3990
 Db 363 TACACAAGCGTGAAGACTTTGGGTTGTGTGACAGCTTTTCTTCCAAACACACTACCC 304
 QY 3991 CCCTGAG 4050
 Db 303 CCCTGAG 244
 QY 4051 TCAGACCGGGGAGATGCGGATGAGCATGAGCACTGTGAACAACATGTGTAACAGAGT 4110
 Db 243 TCAGACCGGGGAGATGCGGATGAGCATGAGCACTGTGAACAACATGTGTAACAGAGT 184
 QY 4111 GGCGGAGAGACTACCTCAACCACTTACCCAGCCGAGCCAACTCAAGTGCCTCTT 4170
 Db 183 GGCGGAGAGACTACCTCAACCACTTACCCAGCCGAGCCAACTCAAGTGCCTCTT 124
 QY 4171 CCTGAGAGCTTACTCTTCAACCTCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4230
 Db 123 CCTGAGAGCTTACTCTTCAACCTCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 64
 QY 4231 GAAGCCCGGAGGTGCTTACTGAGGCTGTCCAGTGGAGCGGAGTGGGCTTGTGTG 4290

QY 1296 GAACATGGGACCGTTACCAACCTGGGAAATCTCCGGAAATTCACCTTCCCTGAA 1355
 DB 437 GAACATGGGACCGTTACCAACCTGGGAAATCTCCGGAAATTCACCTTCCCTGAA 496
 QY 1356 GGGCTTCTCTGTTGGCACTGGGAAAGAACCAAGTCTTAATGCTTTTAAACAGGCTGT 1415
 DB 497 GGGCTTCTCTGTTGGCACTGGGAAAGAACCAAGTCTTAATGCTTTTAAACAGGCTGT 556
 QY 1416 GGGAGAGGCGGAGCTGAGATCTACCTGTCAGGCGGAGGAGGCTGTGATGAA 1475
 DB 557 GGGAGAGGCGGAGCTGAGATCTACCTGTCAGGCGGAGGAGGCTGTGATGAA 616
 QY 1476 GAAAGACAGAGATACCTTTGAGAGAGCTGGAAGATTAATACCTGTTTATAGGCGG 1535
 DB 617 GAAAGACAGAGATACCTTTGAGAGAGCTGGAAGATTAATACCTGTTTATAGGCGG 676
 QY 1536 CAATGACCTCTGATGATTTGCAATGATCTGTCCTACTTATTTCCCAAGACTTCACAGA 1595
 DB 677 CAATGACCTCTGATGATTTGCAATGATCTGTCCTACTTATTTCCCAAGACTTCACAGA 736
 QY 1596 CAATGAGAAAGCCCTGAGACATCTCCAGTCT-GAGGTTCTCCGGGCAATTTG-TGAC 1653
 DB 737 CAATGAGAAAGCCCTGAGACATCTCCAGTCTCCAGTCTCCGCGCAATTTGTTGACC 796
 QY 1654 CTGGTGAAGGCTG--CTTGAGATCGT-CAACCTGAGGAGCTGTACAGAGAAATAA 1707
 DB 797 CTGGTGAAGGCGGCTTTGAGATCTCTCAACCTGAGGAGCTGTACAGAGAAATAA 853

RESULT 9
 LOCUS B1912651 880 bp mRNA linear EST 16-OCT-2001
 DEFINITION 603290988F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:522033 5',
 mRNA sequence.
 VERSION B1912651
 KEYWORDS B1912651.1 GI:16176846
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: c9abs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 http://image.llnl.gov
 plate: LLAM1553 row: 0 column: 22
 High quality sequence start: 27
 High quality sequence stop: 726.
 Location/Qualifiers
 1. 880
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_image="522033"
 /issue_type="leukocyte"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORE6, Site 1: NotI; Site 2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for

FEATURES
 source

full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH MGC Library."

BASE COUNT 201 a 271 c 228 g 179 t 1 others
 ORIGIN
 Query Match 16.2%; Score 708.2; DB 12; Length 880;
 Best Local Similarity 98.3%; Pred. No. 5.3e-173;
 Matches 768; Conservative 0; Mismatches 8; Indels 5; Gaps 5;
 QY 2619 TGCAGCCAACTTGTGACCATCTCCGCAATGCTTTGAGCTCTGATGAGAGTGCC 2678
 DB 27 TGCAGCCAACTTGTGACCATCTCCGCAATGCTTTGAGCTCTGATGAGAGTGCC 86
 QY 2679 CAGAGTCTGGTCAACTGCTGAGCTTCCGAAGCCCAATGATGAGAGTGCTTCT 2738
 DB 87 CAGAGTCTGGTCAACTGCTGAGCTTCCGAAGCCCAATGATGAGAGTGCTTCT 146
 QY 2739 GGGAAACCCAGACAGATGCTCCAGTGCAGAGGCGGTTTGTGTAATGCTGTGAC 2798
 DB 147 GGGAAACCCAGACAGATGCTCCAGTGCAGAGGCGGTTTGTGTAATGCTGTGAC 206
 QY 2799 CTGCG-GGAGAACTCCCAAGAGCTAGCCAGGCTGAGAGCTTCAAGCCTTACCGGA 2857
 DB 207 CTGCGCGGAGAACTCCCAAGAGCTAGCCAGGCTTCAAGCCTTACCGGA 266
 QY 2858 GCAGCATGGCGGAGCTGTGGGGTCAAGGCGCTATGACAGAGAGAGATTTCTGTGG 2917
 DB 267 GCAGCATGGCGGAGCTGTGGGGTCAAGGCGCTATGACAGAGAGAGATTTCTGTGG 326
 QY 2918 TGCTGAGCCCTTCTCCAGAACATCAGCTCCCTGCTGCGAGATGGGCTCCAGATA 2977
 DB 327 TGCTGAGCCCTTCTTCCAGAACATCAGCTCCCTGCTGCGAGATGGGCTCCAGATA 386
 QY 2978 GGTCTTCTTTTGGCCAGACTGATCCACCAATGAGAAATTCATCTCCAGCTGGCA 3037
 DB 387 GGTCTTCTTTTGGCCAGACTGATCCACCAATGAGAAATTCATCTCCAGCTGGCA 446
 QY 3038 GAGCCCTTTGAGAACATATGCTTGAACCACTTGGAGCAAAAGAGACCTTGAGCTGA 3097
 DB 447 GAGCCCTTTGAGAACATATGCTTGAACCACTTGGAGCAAAAGAGACCTTGAGCTGA 506
 QY 3098 GAGCAGAGATGCCATCACTGCTCCCACTCAGAAATGAGCCCTTCTGAGAAACCTTCGGA 3157
 DB 507 GAGCAGAGATGCCATCACTGCTCCCACTCAGAAATGAGCCCTTCTGAGAAACCTTCGGA 566
 QY 3158 ATAGTAACATACAGTACCCCATCAAGCCAGCAATGAGAACTGGGAGAGTCTCTGT 3217
 DB 567 ATAGTAACATACAGTACCCCATCAAGCCAGCAATGAGAACTGGGAGAGTCTCTGT 626
 QY 3218 GTACAGAGTGAAGGCTTCAATAGTTCGAAC-CTGTGTCACAGAGTTCGACAGCA 3276
 DB 627 GTACAGAGTGAAGGCTTCAATAGTTCGAAC-CTGTGTCACAGAGTTCGACAGCA 686
 QY 3277 GACATCAAGTGTGGCGGCGGCTGGTGACTCTGACTACCA-GCAGTGGAGAGTGGAGC 3335
 DB 687 GACATCAAGTGTGGCGGCGGCTGGTGACTCTGACTACCA-GCAGTGGAGAGTGGAGC 746
 QY 3336 AAACAATCTCAGTACCTTCCATATGCTTGG-AGGGAGACTCT-CTTGAAGATGAGAGG 3393
 DB 747 CAGCAATCTCAGTACCTTCCATATGCTTGG-AGGGAGACTCT-CTTGAAGATGAGAGG 806
 QY 3394 G 3394
 DB 807 G 807
 RESULT 10
 LOCUS BX354650 1186 bp mRNA linear EST 05-MAY-2003
 DEFINITION BX354650 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 cDNA clone CS0DC03YA08 5-PRIME, mRNA sequence.
 ACCESSION BX354650

```

VERSION      BX34650.1  GI:30379890
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE    1 (bases 1 to 1186)
AUTHORS      L.M.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished
COMMENT      Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 EVRY cedex - France
              Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
              Library was constructed by Life Technologies, a division of
              Invitrogen. This sequence belongs to sequence cluster 1474.f,
              Contact : Feng Liang Email : fliang@lifetech.com URL :
              http://fulllength.invitrogen.com/Invitrogen Corporation 1600
              Faraday Avenue Genoscope sequence ID : CS0DC023BA04Qp1.
FEATURES
  source
    1..1186
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0DC023YA08"
      /issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
      /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
      /note="1st strand cDNA was primed with a NotI-oligo(dT)
      primer. Five prime end enriched, double-strand cDNA was
      digested with Not I and cloned into the Not I and EcoR V
      sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT   288 a 315 c 281 g 202 t 100 others
ORIGIN
Query Match 15.7%; Score 687.8; DB 13; Length 1186;
Best Local Similarity 91.0%; Pred. No. 1.4e-167;
Matches 711; Conservative 25; Mismatches 42; Indels 3; Gaps 1;

QY 2929 TTCTTCAGAAATCCAGCTCCCTGCTGCGGATGGGCTCCAGATACCTCTTCTT 2968
    |||||
DB 65 TTCTTCAGAAATCCAGCTCCCTGCTGCGGATGGGCTCCAGATACCTCTTCTT 124

QY 2989 GCCCAGACTGATCCACCCCAATTCAGAAATTCACCTCCAGCTGGCAGAGCCCTTTGG 3048
    |||||
DB 125 GCCCAGACTGATCCACCCCAATTCAGAAATTCACCTCCAGCTGGCAGAGCCCTTTGG 184

QY 3049 ACCAAATGCTTGAACCACTTGAAGAAAAGAGACCCCTGACCTGAGAGCAGAGATG 3108
    |||||
DB 185 ACCAAATGCTTGAACCACTTGAAGAAAAGAGACCCCTGACCTGAGAGCAGAGATG 244

QY 3109 CCCATCACTGCTCCCACTAGAAATGAGCCCTTCTTGAAACCCCTCGGAATAGTAATAC 3168
    |||||
DB 245 CCCATCACTGCTCCCACTAGAAATGAGCCCTTCTTGAAACCCCTCGGAATAGTAATAC 304

QY 3169 ACGTACCCCAATCAAGCAGCATTGAGAACTGGGGCAGTCTCTGTGACAGAGTGG 3228
    |||||
DB 305 ACGTACCCCAATCAAGCAGCATTGAGAACTGGGGCAGTCTCTGTGACAGAGTGG 364

QY 3229 AAGGCTTCATATAGTTTCCAACTCTGTCCACGAGCTCCGACAGAGACATCAAGTG 3288
    |||||
DB 365 AAGGCTTCATATAGTTTCCAACTCTGTCCACGAGCTCCGACAGAGACATCAAGTG 424

QY 3289 GTGGCCCGCTGGGTGACTCTGTGACTACAGCAGTGGAGTGCACCAAACTCCAGT 3348
    |||||
DB 425 GTGGCCCGCTGGGTGACTCTGTGACTACAGCAGTGGAGTGCACCAAACTCCAGT 484

QY 3349 GACCTACCAACATCTTGGAGGGAGCTCTTGGAGCAATGGAGGGATGGGAATTGGAG 3408
    |||||
DB 485 GACCTACCAACATCTTGGAGGGAGCTCTTGGAGCAATGGAGGGATGGGAATTGGAG 544

QY 3409 ACTGACACCAACATCTGCCAAC--ATTCTGAAGAAGTTCAACCTTACTCTTGGCTTC 3465
    |||||
DB 545 ACTGACACCAACATCTGCCAAC--ATTCTGAAGAAGTTCAACCTTACTCTTGGCTTC 604

```

```

QY 3466 TCTACGACGACTGGAGGGGAGCAGCAGGACTAATGTGGCAGCGGAGGGCCAGAGCT 3525
    |||||
DB 605 TTATCCAGACACTGGAGGGGAGCAGCAGGACTAATGTGGCAGCGGAGGGCCAGAGCT 664

QY 3526 AGGAGACATGCGACCCAGCGCTGGGAGCTGTGTAGAGCGAATGAAAAACAGCCCGACATC 3585
    |||||
DB 665 AGGAGACATGCGACCCAGCGCTGGGAGCTGTGTAGAGCGAATGAAAAACAGCCCGCTATY 724

QY 3586 AACCTGGAGAAAGATCGAAGCTGTGCACCTTTATTTGGGTCAACGCTGTGTGCAT 3645
    |||||
DB 725 AATCTGGAGAAAGATCGAAGCTGTGCACCTTTATTTGGGTCAACGATTTGTGTGCAT 784

QY 3646 TACTGTGAGAAATCGGAGGCGCCACTTGGCCAGGAATATTTGACAGACATCCAAAGGCC 3705
    |||||
DB 785 TACTGTGATATYCTGTGGGCGCCCGACCAACCCCTTTKGAAGCTGAGAGAGAGAGGC 844

QY 3706 C 3706
    |
DB 845 C 845

RESULT 11
BM549919 1033 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT 6544181 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5745887
DEFINITION 5', mRNA sequence.
ACCESSION BM549919
VERSION BM549919.1 GI:18785712
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE    1 (bases 1 to 1033)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgsbds-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LINL at:
              http://image.lnl.gov
              Plate: L1AM12769 row: 1 column: 24
              High quality sequence stop: 626.
FEATURES
  source
    1..1033
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:5745887"
      /issue_type="leukocyte"
      /lab_host="DH10B"
      /clone_lib="NIH_MGC_118"
      /note="Vector: pCMV-SPORT6, Site 1: NotI, Site 2: EcoRV
      (destroyed); RNA source leukocytes from anonymous pool of
      non-activated adult donors. Library is oligo-dT primed
      and directionally cloned (EcoRV site is destroyed upon
      cloning). Average insert size 1.7 kb, insert size range
      1.2-3.3 kb. Library is normalized and enriched for
      full-length clones and was constructed by C. Gruber
      (Invitrogen). Research Genetics tracking code 027. Note:
      this is a NIH MGC Library."
BASE COUNT   261 a 292 c 291 g 188 t 1 others
ORIGIN
Query Match 15.5%; Score 677.4; DB 12; Length 1033;
Best Local Similarity 98.5%; Pred. No. 6.4e-165;
Matches 736; Conservative 0; Mismatches 6; Indels 5; Gaps 5;

```

```

QY 3057 GCTTGAACCACTTGAAGCAAAAAGAGACCCTGAGCCTGAGAGAGAGATGCCATCAC 3116
DB 153 GCTTGAACCACTTGAAGCAAAAAGAGACCCTGAGCCTGAGAGAGAGATGCCATCAC 212
QY 3117 CTGTCCCACTCAGAAATGAGCCCTTCTGAGAACCCCTGGAATATGTAATCAACGTAACC 3176
DB 213 CTGTCCCACTCAGAAATGAGCCCTTCTGAGAACCCCTGGAATATGTAATCAACGTAACC 272
QY 3177 CATCAAGCCAGCCATTTGAGAACTGGGGCAGTGAATCTCTGTATCAGAGTGAAGGCTTC 3236
DB 273 CATCAAGCCAGCCATTTGAGAACTGGGGCAGTGAATCTCTGTATCAGAGTGAAGGCTTC 332
QY 3237 CAATAGTGTTCACCACTCTGTCCACAGCTCCGACAGAGACATCAAGTGTGCGCG 3296
DB 333 CAATAGTGTTCACCACTCTGTCCACAGCTCCGACAGAGACATCAAGTGTGCGCG 392
QY 3297 CTTGGGTGACTCTCTGATCTACAGCAGTGGAGCTCGAACAACTCCAGTGAACCTACC 3356
DB 393 CTTGGGTGACTCTCTGATCTACAGCAGTGGAGCTCGAACAACTCCAGTGAACCTACC 452
QY 3357 CAGATCTTGAAGGGGACTCTCTTGGAGCATTTGAGGGGATGGGAATCTGAGACTCACAC 3416
DB 453 CAGATCTTGAAGGGGACTCTCTTGGAGCATTTGAGGGGATGGGAATCTGAGACTCACAC 512
QY 3417 CACACTGCCCCAATCTTGAAGAACTCAACCTTACCTCTGAGCTTCTCTACAGCAC 3476
DB 513 CACACTGCCCCAATCTTGAAGAACTCAACCTTACCTCTGAGCTTCTCTACAGCAC 572
QY 3477 CTGGGAGGGGACAGCAGAGACTAAATGTGACGGGAAAGGGGACAGACTTACGATGCC 3536
DB 573 CTGGGAGGGGACAGCAGAGACTAAATGTGACGGGAAAGGGGACAGACTTACGATGCC 632
QY 3537 AGCCCAAGGCTTGGGACCTGTGTAGAGCGAATG-AAAAACAGCCCGGACATCAACCTGG-AG 3594
DB 633 AGCCCAAGGCTTGGGACCTGTGTAGAGCGAATG-AAAAACAGCCCGGACATCAACCTGG-AG 692
QY 3595 AAAGACTGGAAGGCTGTCACACTCTTCAATGGGGTCAAGAGCTTGTGTCTTACTGAG 3654
DB 693 AAAGACTGGAAGGCTGTCACACTCTTCAATGGGGTCAAGAGCTTGTGTCTTACTGAG 752
QY 3655 AATCCGAGAGCCCACTTGGCCAGGATATGTTTCAAGACATCCAAACAGGCCCTGAGCATC 3714
DB 753 AATCCGAGAGCCCACTTGGCCAGGATATGTTTCAAGACATCCAAACAGGCCCTGAGCATC 812
QY 3715 CTCTCTGAGAGAGCTCCCA-GGGCTTTCGTCAAGCT-GGTGAGGTGATGAGCT-GGCT 3771
DB 813 CTCTCTGAGAGAGCTCCCA-GGGCTTTCGTCAAGCT-GGTGAGGTGATGAGCT-GGCT 872
QY 3772 AGCTGTACCAAGGCGCAAGGCGGGAAA 3798
DB 873 AGCTGTACCAAGGCGCAAGGCGGGAAA 899

RESULT 12
LOCUS B1912073 763 bp mRNA linear EST 16-OCT-2001
DEFINITION 603068631F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5217606 5',
ACCESSION B1912073
VERSION B1912073.1 GI:16176045
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 763)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@rs-rcmail.nih.gov
Tissue Procurement: Life Technologies, Inc.

```

```

FEATURES
    source
        1..763
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5217606"
            /issue_type="leukocyte"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_118"
            /note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
            (destroyed); RNA source leukocytes from anonymous pool of
            non-activated adult donors. Library is oligo-dT primed
            and directionally cloned (EcoRV site is destroyed upon
            cloning). Average insert size 1.7 kb, insert size range
            1.2-3.3 kb. Library is normalized and enriched for
            full-length clones and was constructed by C. Gruber
            (Invitrogen). Research Genetics tracking code 027. Note:
            this is a NIH MGC Library."
BASE COUNT 177 a 238 c 193 g 155 t
ORIGIN

```

```

Query Match 15.3%; Score 671.4; DB 12; Length 763;
Best Local Similarity 95.9%; Pred. No. 1.9e-163;
Matches 722; Conservative 0; Mismatches 26; Indels 5; Gaps 4;

```

```

QY 2612 TGTATTCTGACGCAACTTTGTTGACATCTCCGCAATGCTTGGAGCTCTGATGAG 2671
DB 2 TGTATTCTGACGCAACTTTGTTGACATCTCCGCAATGCTTGGAGCTCTGATGAG 60
QY 2672 AGGTGCGCAGAGCTCGTGTCAACCTGTGAGACTTCTGAAACCCCACTATCATCGGAGG 2731
DB 61 AGGTGCGCAGAGCTCGTGTCAACCTGTGAGACTTCTGAAACCCCACTATCATCGGAGG 120
QY 2732 TGTTCCTGGGAAACCCAGACAAAGTGCAGAGCAGAGCCGTTTGTGTAATGCG 2791
DB 121 TGTTCCTGGGAAACCCAGACAAAGTGCAGAGCAGAGCCGTTTGTGTAATGCG 179
QY 2792 TTCTGACCTTGGGAGAACTCCCAAGAGCTGAGGCTGAGGCTTTCAGCCGAGCT 2851
DB 180 TTCTGACCTTGGGAGAACTCCCAAGAGCTGAGGCTGAGGCTTTCAGCCGAGCT 239
QY 2852 ACCGAGCAGCATGCGCGAGCTGTGTGGGTGAGGCGGCTATGACAGAGAGGACTTCT 2911
DB 240 ACCGAGCAGCATGCGCGAGCTGTGTGGGTGAGGCGGCTATGACAGAGAGGACTTCT 299
QY 2912 CTGTGTGCTGACGACCTTCTTCAGAACATCCAGCTCCCTGTCTGGCGATGGGCTCC 2971
DB 300 CTGTGTGCTGACGACCTTCTTCAGAACATCCAGCTCCCTGTCTGGCGATGGGCTCC 359
QY 2972 CAGATACGTCCTTCTTGGCCAGACTGATCCCAATCAGAAATTCACCTCCAGC 3031
DB 360 CAGATACGTCCTTCTTGGCCAGACTGATCCCAATCAGAAATTCACCTCCAGC 419
QY 3032 TGGCCAGAGCCCTTGGAGCAAA-TATGCTTGAACCACTTGGAAAGCAAGAACAGACCT 3089
DB 420 TGGCCAGAGCCCTTGGAGCAAAATATGTTGAACCACTTGGAAAGCAAGAACAGACCT 479
QY 3090 GGAAGCTGAGAGAGAGATGCCATCACTGTCCCATCTGAGATGAGCCCTTCTGAGAAC 3149
DB 480 GGAAGCTGAGAGAGAGATGCCATCACTGTCCCATCTGAGATGAGCCCTTCTGAGAAC 539
QY 3150 CCCTGGAAATGTAATCAACGTAACCCATCAAGCAGCCATTTGAGAACTGGGCGAGTGA 3209
DB 540 CCCTGGAAATGTAATCAACGTAACCCATCAAGCAGCCATTTGAGAACTGGGCGAGTGA 599

```

QY 3210 CTTCCTGTGTACAGAGTGAAGGCTTCCAAATAGTGTCCAAAC-CTGTGTCCACAGAGCTCC 3268
 DB 600 CTTCCTGTGTACAGAGTGAAGGCTTCCAAATAGTGTCCAAAC-CTGTGTCCACAGAGCTCC 659
 QY 3269 GACCAGACAGACATTAAGTGTGGCCCGCTGGGTACTCTGTGACTACAGACAGTGGAG 3328
 DB 660 GACACGACAGCTCAAAAGTGTGGCCCGCTGGGTACTCTGTGACTACAGACAGTGGAG 719
 QY 3329 CTGACCAACCAACTCCAGTACCTACCCACATCTTGGAGGG 3371
 DB 720 CTGACCAACCAACTCCAGTACCTACCCACATCTTGGAGGG 762
 RESULT 13
 BI488437 823 bp mRNA linear EST 28-AUG-2001
 LOCUS 603020936F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5191712 5'
 DEFINITION mRNA sequence.
 ACCESSION BI488437
 VERSION BI488437.1 GI:15327665
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 823)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 plate: L1AM1479 row: 9 column: 09
 High quality sequence, stop: 808.
 Location/Qualifiers
 1. 823
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5191712"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_114"
 /note="Organ: brain; Vector: pCMV-Sport6; Site: 1: NotI;
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH_MGC Library."
 BASE COUNT 213 a 239 c 214 g 157 t
 ORIGIN
 Query Match 14.3%; Score 627.8; DB 12; Length 823;
 Best Local Similarity 97.2%; Pred. No. 4.7e-152;
 Matches 660; Conservative 0; Mismatches 17; Indels 2; Gaps 2;

DB 266 CATCAAGCCAGCCATTGAGAACTGGGGAGAGTGACTTCTGTGTACAGAGTGAAGGCTTC 325
 QY 3237 CAATAGTGTTCACACTCTGTGTCCACAGAGTCCGACAGAGATCAATAGTGTGGCCGC 3296
 DB 326 CAATAGTGTTCACACTCTGTGTCCACAGAGTCCGACAGAGATCAATAGTGTGGCCGC 385
 QY 3297 CTTGGTGAATCTCTGACTACAGAGTGGAGCTGACCAAACTCCAGTACCTAACC 3356
 DB 386 CTTGGTGAATCTCTGACTACAGAGTGGAGCTGACCAAACTCCAGTACCTAACC 445
 QY 3357 CACATCTTGGAGGGAGCTCTCTTGGAGCATTTGAGGGAGTGGAACTTGGAGACTCA 3416
 DB 446 CACATCTTGGAGGGAGCTCTCTTGGAGCATTTGAGGGAGTGGAACTTGGAGACTCA 505
 QY 3417 CACACTGCCCAATTTCTGAAGAAGTTCAACCTTACTCTCTTGGCTTCTTACAGAC 3476
 DB 506 CACACTGCCCAATTTCTGAAGAAGTTCAACCTTACTCTCTTGGCTTCTTACAGAC 565
 QY 3477 CTGGAGAGGGAGACAGAGACTTAATGTGGAGGGAGGGAGGAGCTAGGAGCATGCC 3536
 DB 566 CTGGAGAGGGAGACAGAGACTTAATGTGGAGGGAGGGAGGAGCTAGGAGCATGCC 625
 QY 3537 AGCCAGGCTCTGGAGCTGTGAGGCAATGAAGAAACAGCCCGACATCAACCTGAGAA 3596
 DB 626 AGCCAGGCTCTGGAGCTGTGAGGCAATGAAGAAACAGCCCGACATCAACCTGAGAA 685
 QY 3597 AGACTGAGAGCTGTGACACTTTC-ATTGGGGTCAACGACTTGTCTATTACTGTGAGA 3655
 DB 686 AGACTGAGAGCTGTGACACTTTCATTGGGGTCAACGACTTGTCTATTACTGTGAGA 745
 QY 3656 ATCCGAGGGCCACTTGGCCAGGAAATATGTTCAGACATCCAAAGGCTTGGACATGC 3715
 DB 746 ATCCGAGGGCCACTTGGCCAGGAAATATGTTCAGACATCCAAAGGCTTGGACATGC 804
 QY 3716 TCTCTGAGAGCTGCCAAG 3724
 DB 805 TCTCTGAGAGCTGCCAAG 823

RESULT 14
 BM714799 607 bp mRNA linear EST 28-FEB-2002
 LOCUS BM714799
 DEFINITION U1-E-EJ0-ahs-f-21-0-U1.r1 U1-E-EJ0 Homo sapiens cDNA clone
 U1-E-EJ0-ahs-f-21-0-U1 5', mRNA sequence.
 ACCESSION BM714799
 VERSION BM714799.1 GI:19028057
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 607)
 Bonaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 JOURNAL MEDLINE
 PUBMED 8889548
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: Bento-Soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES
 source
 Location/Qualifiers
 1.607
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-EJ0-ahs-f-21-0-UI"
 /tissue_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"
 /dev_stage="fetal and adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-EJ0"
 /note="Organ: eye; Vector: pRT73-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I; UI-E-EJ0 is a subcloned cDNA library constructed according to Bonaldo, Lemmon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pRT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dri)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTGT; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCCA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."
 BASE COUNT 156 a 178 c 156 g 117 t
 ORIGIN
 Query Match 13.9%; Score 607; DB 12; Length 607;
 Best Local Similarity 100.0%; Pred. No. 9.9e-147;
 Matches 607; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 541 GAGATGGCCATCGCACTCTGGAACACATGCTGGAACCATGGGCGGCAAGACTACTCC 600
 |||
 Qy 4129 AACACT 4135
 |||
 Db 601 AACACT 607
 |||
 RESULT 15
 AK015993
 LOCUS
 DEFINITION
 Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4930539A06 product:weakly similar to PHOSPHOLIPASE B [Rattus norvegicus], full insert sequence.
 AK015993
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 AK015993.1 GI:12854560
 HTC, CAP trapper.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159
 3 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iizawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system—384-format sequencing pipeline with 384 multichannel sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861
 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S., Aizawa, K., Iizawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Aeburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nishida, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderrelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carinci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Guelinich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Maehima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schobach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Wetz, C., Whitaker, C., Wilming, L., Wymshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohseki, S. and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 685-690 (2001)
 21085660
 11217851
 5
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE


```
QY 1033 AGAAACGCACTACTGACCGAGCTGCAGAAACCCCAAGCAAGCTTGAAGTAAAGAA 1092
    |||||
Db 1236 AAAAATAGCAACTACGAGCCAGACGACTGAACCTATTAACAAGCTTCAGATGAAGAA 1295
    |||||
QY 1093 GGACCGGAATTCAGATGTCTGACAAAGACCCCTCCGATAGGGTTCCCACTCAATTCAAT 1152
    |||||
Db 1296 GGATCAGAAATTCACCTGTCTGACAGAACCCCTCGAACTCAATCCCAACACAGGTACA 1355
    |||||
QY 1153 AGGCTG 1158
    |||||
Db 1356 TGGCTG 1361
```

Search completed: January 6, 2004, 14:41:56
Job time : 8450 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 6, 2004, 18:56:30 ; Search time 27 Seconds
(Without alignments)
2539.442 Million cell updates/sec

Title: US-10-054-691-2
Sequence: 1 MGIRPGIFLELLILGQGT.....RCRGRGRDPPMSLRVAL 1458

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5987.5	77.1	1458	1	PHLX_RABIT
2	149	1.9	1458	1	YAMB_SCHPO
3	138	1.8	1455	1	FACA_HUMAN
4	131.5	1.7	1361	1	GLI4_XENLA
5	128	1.6	1155	1	C2TA_MOUSE
6	126.5	1.6	1534	1	YFAS_ECOLI
7	126.5	1.6	6885	1	SNE2_HUMAN
8	125.5	1.6	1178	1	ATC1_YARLI
9	125	1.6	1178	1	YFAS_ECOLI
10	123.5	1.6	1534	1	YFAS_ECOLI
11	123.5	1.6	2672	1	GCN1_YEAST
12	123	1.6	1039	1	PD2P_ARATH
13	122.5	1.6	713	1	GALK_SCHPO
14	122	1.6	1811	1	SC71_SCHPO
15	121.5	1.6	1097	1	TOLR_DROME
16	121.5	1.6	8797	1	SNE1_HUMAN
17	121	1.6	953	1	ATC3_BOVIN
18	121	1.6	4725	1	DYHC_DICDI
19	120.5	1.6	2843	1	APC_HUMAN
20	119.5	1.5	1192	1	RTN4_HUMAN
21	119.5	1.5	1276	1	BXD_CLOBO
22	117.5	1.5	1279	1	BCSC_PSEFL
23	116.5	1.5	1302	1	UBPY_CAEEL
24	116.5	1.5	2167	1	BEM2_YEAST
25	116.5	1.5	2469	1	TEGU_HVISA
26	116	1.5	919	1	ATC1_HUMAN
27	116	1.5	1420	1	SRB3_YEAST
28	116	1.5	1489	1	CC25_YEAST
29	115.5	1.5	555	1	G6P1_BUCAP
30	115.5	1.5	1849	1	BIG1_HUMAN
31	115.5	1.5	5596	1	MDN1_HUMAN
32	115	1.5	1301	1	SAC3_YEAST
33	114.5	1.5	1498	1	YIA9_CLOAB

34	114.5	1.5	2363	1	SPOC_MOUSE
35	114	1.5	919	1	ATC1_RAT
36	114	1.5	2738	1	PGCV_RAT
37	114	1.5	4377	1	ANK3_HUMAN
38	113.5	1.5	880	1	IMB2_HUMAN
39	113.5	1.5	1770	1	R115_YEAST
40	113	1.5	879	1	YE07_MYCPN
41	113	1.5	954	1	DRP2_HUMAN
42	113	1.5	1489	1	RNG2_SCHPO
43	113	1.5	1533	1	Y566_HUMAN
44	113	1.5	1932	1	PAB1_SCHPO
45	112.5	1.4	882	1	DIS1_SCHPO

ALIGNMENTS

RESULT 1	ID	PHLX_RABIT	STANDARD;	PRT; 1458 AA.
AC	005017;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Phospholipase A2B-precursor (EC 3.1.1.-)			
OS	Oryctolagus cuniculus (Rabbit)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_TaxID=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RC	TISSUE=Intestine;			
RX	MEDLINE=93286138; PubMed=8509424;			
RA	Boll W., Schmid-Chanda T., Semenza G., Mantel N.;			
RT	Messenger RNA expressed in intestine of adult but not baby rabbits.			
RT	Isolation of cognate cDNAs and characterization of a novel brush			
RT	border protein with esterase and phospholipase activity."			
RL	J. Biol. Chem. 268:12901-12911(1993)			
CC	- FUNCTION: HAS ESTERASE AND PHOSPHOLIPASE A/LYSOPHOSPHOLIPASE			
CC	ACTIVITY. CAN CONVERT PHOSPHATIDYLCHOLINE TO FATTY ACIDS AND			
CC	GLYCEROPHOSPHOCHOLINE. COULD BE INVOLVED IN UPTAKE OF DIETARY			
CC	LIPIDS, POSSIBLY INCLUDING LONG CHAIN RETINYL ESTERS.			
CC	- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. BRUSH BORDER.			
CC	- TISSUE SPECIFICITY: INTESTINE.			
CC	- DEVELOPMENTAL STAGE: EXPRESSED IN THE INTESTINE OF ADULT BUT NOT			
CC	BABY RABBITS.			
CC	- SIMILARITY: BELONGS TO THE "GDSL" FAMILY OF LIPOLYTIC ENZYMES.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
DR	EMBL; Z12841; CAI78303.1; -			
DR	PIR; A45665; A45665.			
DR	InterPro; IPR001087; Lipase_GDSL.			
DR	Pfam; PF00657; Lipase_GDSL; 3.			
DR	PROSITE; PS01098; LIPASE_GDSL_SER; 2.			
KW	Hydrolase; Repeat; Signal; Transmembrane.			
FT	SIGNAL	1	19	POTENTIAL.
FT	CHAIN	20	1458	PHOSPHOLIPASE A2B-precursor.
FT	DOMAIN	20	1415	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1416	1439	POTENTIAL.
FT	DOMAIN	1440	1458	CYTOSOLIC (POTENTIAL).
FT	DOMAIN	39	1403	4 X 308-326 AA APPROXIMATE REPEATS.
FT	REPEAT	39	347	1.
FT	REPEAT	362	707	2.
FT	REPEAT	708	1054	3.
FT	REPEAT	1064	1403	4.
FT	ACT_SITE	400	400	BY SIMILARITY.

FT ACT SITE 747 747 BY SIMILARITY.
 FT ACT SITE 1103 1103 BY SIMILARITY.
 SQ SEQUENCE 1458 AA; 161343 MW; 901608PF47B1062E6 CRC64;
 Query Match 77.1%; Score 5987.5; DB 1; Length 1458;
 Best Local Similarity 77.6%; Pred. No. 0;
 Matches 1130; Conservative 127; Mismatches 197; Indels 3; Gaps 3;

QY 1 MGLRPGFLELLLLGCGTPOHTSPKSTLSEGLMPETLKSPFCNPNKLCVMMPSK 60
 DB 1 MALMPSFELGLPLRGRGADQIQTSNGKTLSEGLMPESLKTFFPCDPTLAESVPS 60
 QY 61 SVSLKSPDKIFVVAIGNLEIPDPGCTGDEKODMTE-RPOVOMGMVTLSDIRFSP 119
 DB 61 SVSLKSPDKIFVVAIGNLEIPDPGCTGDEKODMTE-RPOVOMGMVTLSDIRFSP 120
 QY 120 SVMPVCHTKRVI PHDAEDLMIQAQELVNNKENIQLDFQDKLINVFSNAGCYL 179
 DB 121 SALMPLC-PETRLVPRGGAEDLMWQATELVRSKRENQLDFEHDKLINVFSTVSCFP 179
 QY 180 CPBAQONGLAGVDELMTGLDYLOQGVPRAPVNLVLSVAEVSROYHTWLSAPBPC 239
 DB 180 CPBAQONGLAGVDELMTGLDYLOQGVPRAPVNLVLSVAEVSROYHTWLSAPBPC 239
 QY 240 NCSEETRLAKVVMQSYQEAANSLASRYSEQSEFTVVFQPFYETPSTLSEDPRLQ 239
 DB 240 NCSEETRLAKVVMQSYQEAANSLASRYSEQSEFTVVFQPFYETPSTLSEDPRLQ 239
 QY 240 RCLRETSQTLKVLQNSYLEAMDSILASSKYNQESFAVVFQPFYESSLSALLAEPRLQ 239
 DB 240 RCLRETSQTLKVLQNSYLEAMDSILASSKYNQESFAVVFQPFYESSLSALLAEPRLQ 239
 QY 300 DSTTLLMHLNRMMEPAKGEDEPLSVKGRPMKCPSESEYLPFSYNSNYLTRLOKPDK 359
 DB 300 DSTTLLMHLNRMMEPAKGEDEPLSVKGRPMKCPSESEYLPFSYNSNYLTRLOKPDK 359
 QY 360 LEVREGAETRCPDSDPTSVHRLKPADINVIQGLSTRAGNAGSTPGNVLDVT 419
 DB 360 LEVREGAETRCPDSDPTSVHRLKPADINVIQGLSTRAGNAGSTPGNVLDVT 419
 QY 420 QYRGLSVSGDENITGVTTLANIIRFNPSSLKGFSTGKETSPNAPLNOAVAGRAED 479
 DB 420 QYRGLSVSGDENITGVTTLANIIRFNPSSLKGFSTGKETSPNAPLNOAVAGRAED 479
 QY 480 LPAQARLVLMKNDTRIHFQEDWKIITLFIGNDLCDPNDLVHVSPOFTDNIGKALD 539
 DB 480 LPAQARLVLMKNDTRIHFQEDWKIITLFIGNDLCDPNDLVHVSPOFTDNIGKALD 539
 QY 540 ILHAEPRAVNIWTVLEIYNLRELVOEKVYCPRLMRLSLCPVULKPDNSTELALIE 599
 DB 540 ILHAEPRAVNIWTVLEIYNLRELVOEKVYCPRLMRLSLCPVULKPDNSTELALIE 599
 QY 600 FNKKFOEKTQLIESGRYDREDFTVVVOPFEFENVDMPKTSEGLPDNSFPAPDCHFRSSK 659
 DB 600 FNKKFOEKTQLIESGRYDREDFTVVVOPFEFENVDMPKTSEGLPDNSFPAPDCHFRSSK 659
 QY 660 SHSRASALMNNMLEPVQKTTTHKFNKINITCPNVOVFLRTYKNSMOGHGTLWLP 719
 DB 660 SHSRASALMNNMLEPVQKTTTHKFNKINITCPNVOVFLRTYKNSMOGHGTLWLP 719
 QY 720 RAASALHPTSVHMLRPADIDVVAALGSLTRAGNIGSKRPDDLPVTVYRGLSYSGAGDG 779
 DB 720 RAASALHPTSVHMLRPADIDVVAALGSLTRAGNIGSKRPDDLPVTVYRGLSYSGAGDG 779
 QY 780 RSPSASAPTSVHALRPADIDVVAALGSLTRAGNIGSKRPDDLPVTVYRGLSYSGAGDG 779
 DB 780 RSPSASAPTSVHALRPADIDVVAALGSLTRAGNIGSKRPDDLPVTVYRGLSYSGAGDG 779
 QY 840 DHRHVNFEHDMKVIYVLIGSDLCDYCTDSNLSAANFVHDLRNALDVLHREYPRVAVNL 899
 DB 840 DHRHVNFEHDMKVIYVLIGSDLCDYCTDSNLSAANFVHDLRNALDVLHREYPRVAVNL 899
 QY 900 VDLNPLTQMQLVGNPKCPVQOASVLCVLTLRNSOGLARLEAFSAIYNSMRELY 959
 DB 900 VDLNPLTQMQLVGNPKCPVQOASVLCVLTLRNSOGLARLEAFSAIYNSMRELY 959
 QY 960 GSGRYDPTQEDFSVVLQPFQNIQLPVVLADGLPDTSPFAPOCIHPNQKFSGLRALMTNM 1019
 DB 960 GSGRYDPTQEDFSVVLQPFQNIQLPVVLADGLPDTSPFAPOCIHPNQKFSGLRALMTNM 1019

DB 960 ESGRYDPTQEDFSVVLQPFQNIQLPVVLADGLPDTSPFAPOCIHPNQKFSGLRALMTNM 1019
 QY 1020 LEPGLSKTETLDIRAEMPITTCPONEPFLTPKNSYITPIKAIENWGSDFLCTEKAS 1079
 DB 1020 LEPGLSKTETLDIRAEMPITTCPONEPFLTPKNSYITPIKAIENWGSDFLCTEKAS 1079
 QY 1080 NSVPTSVHOLRPAIDIKVVAALGSLTTAVGARPNNSSDPTSRGHSWSIGCGDNLETHT 1139
 DB 1080 NSVPTSVHOLRPAIDIKVVAALGSLTTAVGARPNNSSDPTSRGHSWSIGCGDNLETHT 1139
 QY 1140 TLPNILKKNPVLGEGSTSTWEGTAGLNYAEGARAPDPAQAMDVERKNSPDINLEK 1199
 DB 1140 TLPNILKKNPVLGEGSTSTWEGTAGLNYAEGARAPDPAQAMDVERKNSPDINLEK 1199
 QY 1200 DMKLVTLFVGNDLCHFCENPBGSSGEVYQHQAQALDVLYEELPRTFVVNVVEMELAGL 1259
 DB 1200 DMKLVTLFVGNDLCHFCENPBGSSGEVYQHQAQALDVLYEELPRTFVVNVVEMELAGL 1259
 QY 1260 YOGGKGKCA-MLAONNCTCLRHSSSLEKQELKKNVNLQHGISFSFYHGYTOREDFA 1318
 DB 1260 YOGGKGKCA-MLAONNCTCLRHSSSLEKQELKKNVNLQHGISFSFYHGYTOREDFA 1318
 QY 1319 VVVQPFQNTLTPLNREGDTLTFPSEDCPHFSDRGHAEMALMNNMLEPVGRKTTSN 1378
 DB 1319 VVVQPFQNTLTPLNREGDTLTFPSEDCPHFSDRGHAEMALMNNMLEPVGRKTTSN 1378
 QY 1379 FTHSRKALKCPSPESPLYTLTRNSRLPPOAEAPVLYVAVVAAGVGLVIGTIVM 1438
 DB 1379 FTHSRKALKCPSPESPLYTLTRNSRLPPOAEAPVLYVAVVAAGVGLVIGTIVM 1438
 QY 1439 RCRGRGREDPPMSLRT 1455
 DB 1440 RCRGRGREDPPMSLRT 1455

RESULT 2
 YAMB_SCHPO
 ID YAMB_SCHPO STANDARD; PRT; 3655 AA.
 AC Q1064;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C1F5.11c in chromosome I.
 GN SPAC1F5.11C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxId=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Scouras J., Peat N., Hayes J., Baker S., Baaham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Frazer A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitz E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Rabinowitz E.,
 RA Welter I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gebel C., Fuchs M., Filicz C., Holzer B., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Beyer P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerretti L., Lowe T., McCombe W.R., Paulsen I., Potashkin J., Shpakovski G.V., Usery D., Barrell B.G., Nurse P., "The genome sequence of Schizosaccharomyces pombe.", Nature 415:871-880(2002).

CC -1- SIMILARITY: STRONG, TO YEAST YHR099W.
CC -1- SIMILARITY: BELONGS TO THE PI3/P14-KINASES FAMILY.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements or send an email to license@ebi.ac.uk).

CC EMBL: Z68136; CAA22239.1; -.
CC PIR: T38084; T38084.
CC GeneDB SPombe; SPAC1F5.11c; -.
CC InterPro: IPR003151; PAT.
CC InterPro: IPR003152; PATC.
CC InterPro: IPR004043; PI3_P14_kinase.
CC InterPro: IPR001440; TPR.
CC Pfam: PF02259; FAT; 1.
CC Pfam: PF02260; FATC; 1.
CC Pfam: PF00454; PI3_P14_kinase; 1.
CC SMART: SM00146; PI3Kc; 1.
CC PROSITE: PS50290; PI3_4_KINASE_3; 1.
CC Hypothetical protein; Transferrase; Kinase.
CC DOMAIN 3324 3655 PI3K/PI4K (BY SIMILARITY).
CC SEQUENCE 3655 AA; 420774 MW; 50475E3F3FC124A CRC64;

Query Match 1.9%; Score 149; DB 1; Length 3655;
Best Local Similarity 20.0%; Pred. No. 0.76;
Matches 167; Conservative 112; Mismatches 307; Indels 248; Gaps 40;

QY MLEPVGQKTR--HKFNKINITCNOVQPLRTKNSM-----OGHWLQ 717
DB LKPLAFLADLHHVRDELNT--QIRKSMITSTNHDLTLTIGLOTMKARLLLM 423
QY RDR--APSALHPTSHALRPADIOVVAAGSLTANGIGSKPDDLPTVTOYRGLSYS 774
DB VDRMISLPsi-----PDALFILLSIFDSFV-----NKFSLLNSLDGFFKKX 467
QY AGDGSLENTTLLPILREPNRLTGVANGTGANDTNAFLNQAADIMSOVQL 834
DB -----EIKETKSPTRSSPRDLSFSFVSND-----GSFL-----FKNIMFGRLAL 508
QY MCKMKDHRVNFHEDKIVTLVIGSDLDQYCTDSNLYSANPVDLRNALDVLHREVR 894
DB MGLR-----TKSRCLIEGGQSGFGLTNIKPFELVVF-----OKLFEVKK 551
QY VLVNVDVFLNPTIMROVFLGNPDKCFVQQAASVLCNVLTLRENSQELARLEAFSRRSS 954
DB 552 GF-----SYRPF--EQVILETFEC--EESLDRAPISTLPRKDKDCLVYATIF---- 599
QY 955 MEVLGSGRYDQEDFSVVLQPFQNIQLPVLADGLPD--TSFADDCIHPNQ---KTH 1008
DB 600 -----THLEPSIFLKVETN--LPTFFDQKKMLTLPHIQFLSNSTSSKFL 646
QY 1009 SOLARALMTNMLEPFGSKTEITDLRAEMPI-----TCFQNPFFLR----- 1049
DB 647 NTLRLPL-LSRIEELSS--DIRGSVLRLFRLSFTVSMFATENEPEVLRPYSEIT 701
QY 1050 -----TPRNSNTYPIKPAIENMGSDFLCTEKWANSVPTSHOARPADIKVAAL 1100
DB 702 VKCMKLAPNSANSINYYLLRALFRQIGGRFESLYK---EWMPLHALLAENSL-- 755
QY 1101 GSSLTAVAGARPNNSSDLPFS-----WRGLSNISGGDGLTET--- 1137
DB 756 -----ISARTPKERKDLFTLCLTIPVRLSLPYMSYLMRLVMSLXSSQELVAGUR 808

QY 1138 --HTPLNLIKPF-----NPYLGFSTSTW-----EGTAGLVAAEGARADW 1178
DB 809 TELCLCDNLTPPLDIPMAPYIEDLNNALMSHQPLPYNNYNSHTMLKLGKGGNRKL 866
QY 1179 PAQAMDVLRKMSQSPDINEKQWLVTLFGVN--DLCHYCNENPAHLATEYVQIQQAL 1236
DB 869 ---LDRVQSLKNSPEPN--NDFTLILSIRGVQPOLH-----YQYVDEAV 910
QY 1237 DLSEELPRAFAFNVV--WELASLYOGCGCKAMLAONNCTCLHSOSLSKQEL 1291
DB 911 NLUSS--PSSDLEVQQAFTYVCNISKLYVSDATNSLASSIRCAADIKSNFDFRRP 968
QY 1292 KRYNNMLQHGISSFSYWHQYTORDEFAVVQF-----FRONTLTPNBERGDTLTFEFS- 1345
DB 969 YGVIPERMGRSSFT--QLSDSDSTITILASATYGLFATYVDELRE---EAYFWLER 1021
QY 1346 -----DCFHPSD--RGHAEMALALNNMLEPVGRKTSNNFTHSRAKLC 1388
DB 1022 LAVNVIHDIFFAFAVDYIQGNHSGKFTTNLQKEYI-----ISPHYFAHCLSEVVC 1069

RESULT 3
FACA HUMAN STANDARD; PRT; 1455 AA.
ID FACA HUMAN
AC 015360; Q75266; Q92497; Q96118; Q9UEA5; Q9UEB8; Q9UEB9; Q9UPK3;
AC 09Y6M2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Fanconi anemia group A protein (FACA protein).
GN FANCA OR FAA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lymphoblast;
RX MEDLINE=97051928; PubMed=8896563;
RA Lo Ten Foe J.R., Roostmans M.A., Bosnyan-Collins L., Alon N.,
RA Wilker M., Parker L., Lightfoot J., Carreau M., Callen D.F.,
RA Savota A., Cheng N.C., van Berkel C.G.M., Strunk M.H.P., Gille J.J.P.,
RA Pals G., Kruty F.A.E., Pronk J.C., Arwert F., Buchwald M., Jansen H.,
RT "Expression cloning of a cDNA for the major Fanconi anemia gene,
RT FAA".
RL Nat. Genet. 14:320-323(1996).
RN (2)
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=97312685; PubMed=9169126;
RA Ianzano L., D'Agostino M., Centra M., Savino M., Leyran O.,
RA Auerbach A.D., Cleton-Jansen A.M., Doggett N.A., Pronk J.C.,
RA Tipping A.J., Gibson R.A., Mathew C.G., Whitmore S.A., Apostolou S.,
RA Callen F.C., Zelman L., Savola A.,
RT "The genomic organization of the Fanconi anemia group A (FAA) gene.",
RL Genomics 41:309-314(1997).
RN (3)
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Riche D.O., Bruce D., Mundt M., Doggett N., Munk C., Saunders E.,
RA Robinson D., Jones M., Buckingham J., Chasteen L., Thompson S.,
RA Goodwin L., Bryant J., Tesmer J., Melnick L., Longfibre J., White S.,
RA Ueng S., Tatum O., Campbell C., Fawcett J., Maltbie M., Deaven L.,
RT "Sequencing of human Fanconi anemia complementation group A gene
RT genomic region.",
RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
RN (4)
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Cervix;
RX MEDLINE=22388257; PubMed=12477932;
RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buerow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,
RA Brownstein M.J., Udell T.B., Tohyuki S., Cantucci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzyzanski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL [5]
RN SEQUENCE OF 491-571 AND 610-671 FROM N.A.
RX MEDLINE=98389634; PubMed=9721219;
RA Centra M., Memo E., D'Apollito M., Savino M., Ianzano L.,
RA Notarangelo A., Liu J., Doggett N.A., Zelante L., Savio A.,
RT "Fine exon-intron structure of the Fanconi anemia group A (FAA) gene
and characterization of two genomic deletions.";
RL Genomics 51:463-467(1998).
RN [6]
RX SEQUENCE OF 491-542 FROM N.A.
RX MEDLINE=98375680; PubMed=9711872;
RA Levan O., Doggett N.A., Auerbach A.D.,
RT "Identification of Alu-mediated deletions in the Fanconi anemia gene
FAA.";
RL Hum. Mutat. 12:145-152(1998).
RN [7]
RP SUBCELLULAR LOCATION, AND MUTAGENESIS.
RX MEDLINE=98414611; PubMed=9742112;
RA Naef D., Kupfer G.M., Suliman A., Lambert K., D'Andrea A.D.,
RT "Functional activity of the Fanconi anemia protein FAA requires PAC
binding and nuclear localization.";
RL Mol. Cell. Biol. 18:5952-5960(1998).
RN [8]
RP PHOSPHORYLATION.
RX MEDLINE=99007271; PubMed=9789045;
RA Yamashita T., Kupfer G.M., Naef D., Suliman A., Joenje H., Asano S.,
RA D'Andrea A.D.,
RT "The Fanconi anemia pathway requires FAA phosphorylation and FAA/PAC
nuclear accumulation.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:13085-13090(1998).
RN [9]
RP VARIANTS FA, AND VARIANTS.
RX MEDLINE=98058767; PubMed=9371798;
RA Levan O., Erlich T., Magdalena N., Gregory J.J., Batisti S.D.,
RA Verlinder P.C., Auerbach A.D.,
RT "Sequence variation in the Fanconi anemia gene FAA.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:13051-13056(1997).
RN [10]
RX VARIANTS FA, PRO-1110 AND GLY-1117.
RX MEDLINE=99225041; PubMed=10210316;
RA Yamashita T., Naef D., Garcia-Higuera I., Wask J., Cheng A.,
RA Yamashita T., Hipping A., Morgan N., Mathew C.G., D'Andrea A.D.,
RT "A patient-derived mutant form of the Fanconi anemia protein, FANCA,
is defective in nuclear accumulation.";
RL Exp. Hematol. 27:587-593(1999).
RN [11]
RX FUNCTION: DNA repair protein that may operate in a postreplication
repair or a cell cycle checkpoint function. May be involved in
interstrand DNA cross-link repair and in the maintenance of normal
chromosome stability.
RN [12]
RX SUBUNIT: Belongs to a multisubunit complex composed of FANCA,
FANCC, FANCF and FANCG proteins.
RN [13]
RX SUBCELLULAR LOCATION: Nuclear (major) and cytoplasmic (minor).
RN [14]
RX ALTERNATIVE PRODUCTS:
RX Event=Alternative splicing; Named isoforms=2;
RX Name=1;
RX IsoId=O15360-1; Sequence=Displayed;
RX Name=2;

CC IsoId=O15360-2; Sequence=VSP_007039;
CC Note=No experimental confirmation available;
CC PTM: Phosphorylation primarily on serine residues. Phosphorylation
CC is required for the formation of the nuclear complex. Not
CC phosphorylated in cells derived from groups A, B, C, E, F, G, and
CC H.
CC [15]
CC DISBASE: Defects in FANCA are the cause of one of the 8
CC complementation groups of Fanconi anemia (FA) [MIM:227650], an
CC autosomal recessive disorder characterized by progressive
CC pancytopenia, a diverse assortment of congenital malformations,
CC and a predisposition to the development of malignancies. At the
CC cellular level it is associated with hypersensitivity to DNA-
CC damaging agents, chromosomal instability (increased chromosome
CC breakage), and defective DNA repair. FANCA defects account for 65%
CC of FA.
CC [16]
CC DATABASE: NAME=Fanconi Anemia Mutation Database;
CC WWW="http://www.rockefeller.edu/fanconi/mutate/";
CC [17]
CC DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/FA1D102.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X59226; CAB67610.1; -
CC EMBL: 283067; CAB05445.1; JOINED.
CC EMBL: 283068; CAB05445.1; JOINED.
CC EMBL: 283069; CAB05445.1; JOINED.
CC EMBL: 283070; CAB05445.1; JOINED.
CC EMBL: 283071; CAB05445.1; JOINED.
CC EMBL: 283072; CAB05445.1; JOINED.
CC EMBL: 283073; CAB05445.1; JOINED.
CC EMBL: 283074; CAB05445.1; JOINED.
CC EMBL: 283075; CAB05445.1; JOINED.
CC EMBL: 283076; CAB05445.1; JOINED.
CC EMBL: 283077; CAB05445.1; JOINED.
CC EMBL: 283078; CAB05445.1; JOINED.
CC EMBL: 283079; CAB05445.1; JOINED.
CC EMBL: 283080; CAB05445.1; JOINED.
CC EMBL: 283081; CAB05445.1; JOINED.
CC EMBL: 283082; CAB05445.1; JOINED.
CC EMBL: 283083; CAB05445.1; JOINED.
CC EMBL: 283084; CAB05445.1; JOINED.
CC EMBL: 283085; CAB05445.1; JOINED.
CC EMBL: 283086; CAB05445.1; JOINED.
CC EMBL: 283087; CAB05445.1; JOINED.
CC EMBL: 283088; CAB05445.1; JOINED.
CC EMBL: 283089; CAB05445.1; JOINED.
CC EMBL: 283090; CAB05445.1; JOINED.
CC EMBL: 283091; CAB05445.1; JOINED.
CC EMBL: 283092; CAB05445.1; JOINED.
CC EMBL: 283093; CAB05445.1; JOINED.
CC EMBL: 283094; CAB05445.1; JOINED.
CC EMBL: 283095; CAB05445.1; JOINED.
CC EMBL: 283151; CAB05445.1; JOINED.
CC EMBL: AC005360; AAC28751.1; -
CC EMBL: AC005565; AAC3304.1; -
CC EMBL: AC005567; AAC33401.1; -
CC EMBL: BC008879; AAH08979.1; -
CC EMBL: A0225084; CAI12393.1; -
CC EMBL: A0225085; CAI12394.1; -
CC EMBL: AF054569; AAC28331.1; -
CC Genew: HGNC:3582; FANCA.
CC MIM: 607139; -
CC GO: GO:0005737; C:cytoplasm; TAS.
CC GO: GO:0005634; C:nucleus; TAS.
CC GO: GO:0006281; P:DNA repair; TAS.
CC GO: GO:0006461; P:protein complex assembly; TAS.

Matches 249; Conservative 140; Mismatches 498; Indels 379; Gaps 58;

QY 56 NMSKSVHS-----LKPSDIK-----VAIGNLEIPRPGTGDELEKQWTERPOOV----- 103
 Db 10 NSPTLMSISARGLSPEVAHEHKEKRGITGLAPPPPGT-----TPREYCHQMA 59
 QY 104 -----MGWTVLS-----DIIRFSPSPMPVCHTKGRVIPHDAE 139
 Db 60 FLASHSPYELLVQSAAGNTSHHDYLTMPDVSRRSPRVTPRL--SRKALSTISPLS 117
 QY 140 DLMTIQAEVLNRNKENIQLDFQDKILNVPFSNAQCYLCPSAQONGLAAGVD----- 194
 Db 118 DASIDQLMTIRTSPTNSL-----YAYINSSSSASGSYGHLISAGASIPASF 165
 QY 195 -----ELMGVILYLOEVPRAFNVLDSSEVAEVSQYHG--TWLSPAEPCNCSEET 246
 Db 166 PIHKPCLSLAALSO-QRSLSSTSGHTPLHPSPTPARQOQALISANPAPPSNNSSAPDS 224
 QY 247 RLAKVVMQMSYQEAAMNSL--ASSRYSEOSFTVVF-QP----- 282
 Db 225 VLAKVSESAVSSTVAVQVHKRSKVTEBEADSVRFQPDHLTDLEDKDECKOQRE 284
 QY 283 FFEYETPSLHSEDPRLQDSTTLAMHLMNRMEPAGEKDE-----PLSVKH-- 327
 Db 285 HIYETNCHMDGCKSEFPTQDLVHINHDIH--GEKKEFVCRWQDSRQKPKAQYML 342
 QY 328 -----GRPKCPSQ-----ESPILFSYRNN-----YITRL 353
 Db 343 VHMRRHTGKPKHCTFEGCFKAYSRLNKLTHLSHTGKPYVCHDEGCKAFSNASDR 402
 QY 354 QKPODKLEVEGAEL--RCPDKPSDVPTVS-----HRLKPADINVGALGDS 400
 Db 403 AKQONRHSNEKYICVPCCTKRYTDPSSLRKHKVTVHGPBEAHTVKKHNDIIOK--PS 460
 QY 401 LTAGNGAGSTPGVNLDTVQYRGLSMSVGDE--NIGVTTLANTLREFPNSLKGSV 456
 Db 461 LPKENG-----DNEASAKLSGREHSDSVSRQEHCLQTRITKTEDNNMQSSP--GGQS 512
 QY 457 GTGKERSPNALNQAVAG-----GRADLVYQARRLVDMKN-----D 494
 Db 513 SCSEBSPYGNNTNIDSGVDVSLAMQSLDLFELESTSPVDSTVSSWSSGRPATPT 572
 QY 495 TRIHFOE-----DMKITLFLIGNDLCDFCNLDVHYSPONFNDIGKALDILAE 544
 Db 573 QRHSAETGAEREIENDNERFLIYPNATCQ--NTRLPTISANGDVIGVBSVAIN-- 629
 QY 545 VPRAF--VNLVTVLEIVNRELQEKVYCPRMILRLCPVCLKPDNSTELATLIEFN 601
 Db 630 --PAIETLSMNDVTMMQNLNERR-----DSTSTL----- 657
 QY 602 KKFOEKTHOLIEGCRYTREDFTVVVOPFFENDUMKTSF--GLPDNSFPAPOCFHSSK 659
 Db 658 -----SSAYTSRR--SSGISPYFSRRSSSETSOFGGLNNSSSADSDPISTD 703
 QY 660 SHSRAAGALNMNMLEPVGOQKTRHKPFENKINIC-----PNOVOPLRTYKNSMQ 709
 Db 704 ASRRSSASASHSGILPNLINTPRAHYRLKAKYAAATGCPPTPLPNDRIGLRKUSLMD 763
 QY 710 GHGTWLP-----CRDRAPSA-LHPTSVALRAPDIQVVAALGDSLTLAGNGISKP 758
 Db 764 GADFPLEPPRQLPVRRRCDOGGNAGLPMYPHEI--PGNNSRRAS--DVPARTAGIDKP 820
 QY 759 DDLPDVTYQRGSLYSAGDGSLENVTTL--PILREFPNRLTGYAVGTGDAN----- 809
 Db 821 --LPRFSRFH-----SNMSNMTLHPPSLSEBRNGQLQHYTSDGGLHRYVSPR 867
 QY 810 -----DTAFL-----NOAVFGAKAEDLM-----SOVQTLMOQKMDHRVNFHEMDKVIIVLI 857
 Db 868 PPSISEVAMBAISCDAVREGD--DDMLPDDVVQYIRSQARRAEPONLOTES----- 920
 QY 858 GGSDDLCDYCTDSNLYSAA-----NFVDHLRNALDVLHREVPV-----LVNLVDPLNPT 906
 Db 921 -----SPARNLOSTKSFHNNTPTPQPARPAGAYLSRNPALAEICGGGTANNQDNMMPV 972

QY 907 IMQVFLGND--KCPVOASVLCNCLTIRENSQELARLEAFSRAVRSRMRELVSGRY 964
 Db 973 QMNEVSSGTYDVSDLEPKQOFAA---GNLAVVQKQKFAQOYGSNQAPMQRANHIMQOGE 1029
 QY 965 DTQEDFSPVYLQPE-----FQNTQLPVYLDGLDPTSFPAEDCI--HPNKFHSOLA 1012
 Db 1030 SVQRINISVNGQRFNNYLQORQOQWSQCI--VSSDFIQORVYSQSMLSRPMQEGOSQIS 1088
 QY 1013 RALMTMTMBPLSGKT-----ETLDDLAEMPITCPTO-----NPPFLRTPNSTYTP 1059
 Db 1089 PSC--NNMVRPBGVHTHAPSNTLHQRLAVHAGPTGFPANNFSVNODGLHP--NAATYVOP 1146
 QY 1060 IKPAIE 1065
 Db 1147 QKNGLE 1152

RESULT 5
 C27A_MOUSE STANDARD; PRT: 1155 AA.
 AC P79621; 046787, 078036; 078109; Q31115; Q9TPP1;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE MHC class II transactivator (CITRA).
 GN MHC27A OR CITRA OR C27A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBI_TaxID=10090;
 RX [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1; 2 AND 3), TISSUE SPECIFICITY, AND INDUCTION.
 RC STRAIN=BALB/c;
 RA MEDLINE=97327562; PubMed=9184229;
 RX Mhlethaler-Motter A., Otten L.A., Steinle V., Mach B.;
 RT "Expression of MHC class II molecules in different cellular and functional compartments is controlled by differential usage of multiple promoters of the transactivator Citra.";
 RL EMBO J. 16:2851-2860 (1997).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC STRAIN=NOD; TISSUE=Spleen;
 RX MEDLINE=97152412; PubMed=8995190;
 RA Sims T.N., Elliott J.F., Ramasbar V., Denney D.W. Jr., Halloran P.F.;
 RT "Mouse class II transactivator: cDNA sequence and amino acid comparison with the human class II transactivator.";
 RL Immunogenetics 45:220-222 (1997).
 RN [3]
 RP SEQUENCE OF 955-1097 FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Thymus;
 RX MEDLINE=9618886; PubMed=8620527;
 RA Panelli M.C., Wang E., Shen S., Schluter S.F., Bernstein R.M., Hersh E.M., Stoeck A., Gangavalli R., Barber J., Jolly D., Apoptay E.T.;
 RA "Interferon gamma (IFNgamma) gene transfer of an EMT6 tumor that is poorly responsive to IFNgamma stimulation: Increase in tumor immunogenicity is accompanied by induction of a mouse class II transactivator and class II MHC.";
 RL Cancer Immunol. Immunother. 42:99-107 (1996).
 CC -1- FUNCTION: ESSENTIAL FOR TRANSCRIPTIONAL ACTIVITY OF THE HLA CLASS II PROMOTER; ACTIVATION IS VIA THE PROXIMAL PROMOTER. NO DNA BINDING OF IN VITRO TRANSLATED CITRA WAS DETECTED. MAY ACT IN A COACTIVATOR-LIKE FASHION THROUGH PROTEIN-PROTEIN INTERACTIONS BY CONTACTING FACTORS BINDING TO THE PROXIMAL MHC CLASS II PROMOTER, TO ELEMENTS OF THE TRANSCRIPTION MACHINERY, OR BOTH. ALTERNATIVELY IT MAY ACTIVATE HLA CLASS II TRANSCRIPTION BY MODIFYING PROTEINS THAT BIND TO THE MHC CLASS II PROMOTER (BY SIMILARITY).
 CC -2- SUBCELLULAR LOCATION: Nuclear.
 CC -3- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1; Synonyms=1;

```

CC      IsoId=p79621-1; Sequence=Displayed;
CC      Name=2; Synonyms=III;
CC      IsoId=p79621-2; Sequence=VSP_007214, VSP_007215;
CC      Name=3; Synonyms=IV;
CC      IsoId=p79621-3; Sequence=VSP_007216;
CC      -1- TISSUE SPECIFICITY: Isoform 1 is expressed at very high levels in
CC      dendritic cells, at very low levels in spleen and thymus and is
CC      not detected in other tissues. Isoform 2 is detected at high
CC      levels in spleen and tonsil as well as in a number of B-lymphocyte
CC      cell lines, and at very low levels in dendritic cells.
CC      -1- INDUCTION: By interferon-gamma.
CC      -1- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
CC      -1- SIMILARITY: Contains 1 NACHT domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@ib-sib.ch).
CC      -----
DR      EMBL; AF000006; AAB92364.2; -
DR      EMBL; AF000007; AAB92365.1; -
DR      EMBL; AF042158; AAC34366.1; -
DR      EMBL; AF042159; AAC34367.1; -
DR      EMBL; AF100709; AAF06838.1; -
DR      EMBL; AF100710; AAF06839.1; -
DR      EMBL; U06853; AAB48859.1; -
DR      EMBL; U46562; AAB5004.1; -
DR      MGD; MGI:108445; C2ca.
DR      InterPro; IPR001611; LRR.
DR      InterPro; IPR003590; LRR_RNinh_sub.
DR      InterPro; IPR007111; NACHT_NTRase.
DR      Pfam; PF00560; LRR; 1.
DR      SMART; SM00368; LRR_RI; 4.
DR      PROSITE; PS50837; NACHT; 1.
KW      Transcription regulation; Activator; Nuclear protein; ATP-binding;
KW      Leucine-rich repeat; Repeat; Alternative splicing.
FT      REPEAT 175 198 LRR 1.
FT      REPEAT 243 269 LRR 2.
FT      DOMAIN 439 749 NACHT.
FT      REPEAT 979 1000 LRR 3.
FT      REPEAT 1008 1031 LRR 4.
FT      REPEAT 1039 1062 LRR 5.
FT      DOMAIN 129 215 ASP/GLU-RICH (ACIDIC).
FT      NP_BIND 445 452 ATP (POTENTIAL).
FT      VARSPLIC 1 77 Missing (in isoform 2).
FT      VARSPLIC 78 94 /FTID=VSP_007214.
FT      VARSPLIC 94 SLOAFTVGRGTSRDHG -> MRLCLVPGSGSYLPELQ
FT      FT (in isoform 2).
FT      FTID=VSP_007215.
FT      VARSPLIC 1 101 Missing (in isoform 3).
FT      CONFLICT 994 994 /FTID=VSP_007216.
FT      SEQUENCE 1155 AA; 127528 MW; F3FPO5DEBFB0CB71 CRC64;
SQ
Query March 1.6%; Score 128; DB 1; Length 1155;
Beat local similarity 22.9%; Pred. No. 3;
Matches 102; Conservative 47; Mismatches 153; Indels 144; Gaps
QY      39 ETLKNSPPPCPN-KLGVNMP-----SKSVSLKPSDKIFAALGNLEPPDP----- 85
DB      337 ETEDTSPSPCGEPSSIKLKMPFAVAFPHSLD---KTKA-----LQSPRGLVA 387
QY      86 -----GTG-----DLEKQDTERPQOYCMG---VMTVLSDIIRYSPSYMPVC 126
DB      388 VELVVARLERGSSNKSQERELATPWTFR--QLAHGLAEVLQVNSDCRR---PGETQYVA 442
QY      127 HTGK-----RIVPHDGAEDLQAQELVRNKKENLQDPQDMKLINVF----- 171
DB      443 VLGAAGGQSKWARTVSH---TWAGQ-----LQYDF-----VFVYPCGL 481

```

```

OY 172 SNAAGCY-----LCPSAQONGLAAGVDELNGVLDYLOQVPRARVNI.VDISEVAENSG 226
Db 482 DRPGDTYHLRDLCPPLSQ-PLAND--DE---VDIYIRQDPRIVLLIDAELEEAQDL 535
OY 227 YHGTWLSAPAPPCNCSEBETRLAKVVMQWSYQEAWNISLLASRYSEBSFT---VVFQPF 283
Db 536 LHGGCGSLSPFPCSL--RGLLAGIIPQKRLRGCTLLTARRGRLAGLSLGAADIIFE-- 590
OY 284 FYETTPSLHSEDPRLODSTLYLAHLNBNMMEPAGEKDEPLSVKGRPMKPCPSQESPYLFS 343
Db 591 ----VPSFSTK---QAKTYRHRHYFEN--SGTAGNQDKALGLIEGQPLLCL-----S 631
OY 344 YRNSNYTLRLQKPKQDKLEVERGAEIRGCDKXDPSPDVPFVSVHRLKRNADINVGALGDSILTA 403
Db 632 YSHSPVNCRAVCQSLSKALLBEGTEAQLD-----CTLGLYVSLG 671
OY 404 GNGAGSTPGNVLDVLTQYRGLSWAGS 429
Db 672 PAQNSPPG-----ALVELAKLAMELG 693

RESULT 6
YFAS_ECOLI
ID YFAS_ECOLI STANDARD; PRT; 1534 AA.
AC P76464; P76465;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein yfas precursor.
VFAS OR B2227/B2228.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
NCBI_TaxID=562;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655.
RX MEDLINE=9742617; PubMed=9278503;
RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=9734980; PubMed=9205837;
RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
RA Oshima T., Oyama S., Saito N., Sampei G., Sato Y., Sivaundaram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT K-12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113 (1997).
[3]
RP CONCEPTUAL TRANSLATION.
RA Coudert E.;
RL Unpublished observations (MAR-2002).
CC -1- SIMILARITY: BELONGS TO THE UPF0192 FAMILY.
CC POSITION 1368 WAS TRANSLATED AS TRP TO RESTORE THE SIMILARITY WITH
CC THE C-TERMINAL REGION OF OTHER HOMOLOGS.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcement)

```

or send an email to license@sb-sib.ch).

CC or send an email to license@sb-sib.ch).
 DR EMBL: AEO00312; AAC75288.1; ALT INIT.
 DR EMBL: AEO00312; AAC75287.1; ALT INIT.
 DR EMBL: D90854; -; NOT_ANNOTATED_CDS.
 DR EMBL: D90851; -; NOT_ANNOTATED_CDS.
 DR EcoGene: EGI4081; yfas.
 KM Hypothetical protein; signal; Complete proteome.
 FT SIGNAL 1 38 POTENTIAL PROTEIN YFAS.
 FT CHAIN 39 1534 HYPOTHETICAL PROTEIN YFAS.
 SQ SEQUENCE 1534 AA; 169845 MW; 73B71431C51F8B6C CRC64;

Query Match 1.6%; Score 126.5; DB 1; Length 1534;
 Best Local Similarity 18.8%; Pred. No. 5.9;
 Matches 278; Conservative 174; Mismatches 556; Indels 473; Gaps 68;

QY 161 QFWKLVNPFSSNAQCCLPSAQOGLA-----AGVDELMLGVLDYQOE 206
 DB 9 QFWHML-SKFFSGAIAACLSLIVGTGLANNADSLPSSNVAAPAGCTFFLLADSSFSSE 67
 QY 207 VPRAFVNLVLESEAVSROYH-----GTWSPAPCESEETTRLAKVVMQWY-- 257
 DB 68 EAK-----VLEAPGRDYRRYOMEYGYGVDRLYRIPMAFLRQOKLHRTIVQPYLG 122
 QY 258 -----QEAANSLASSRYSEOSFTVVPFPFFETTPSLHSIDPRLQDSTTLAMHLMN 310
 DB 123 DGLNLTLTWLMDMWYKGRVWQRTFSSQSRQNTVALPELQGNALIKSR-----YVQN 178
 QY 311 RMMEPAGEKQEPISVYKGRPMKCPSCSPYLFSTRNSNYLTRLOKPDKLEVRGAEIR 370
 DB 179 NQSPPL-KKYPVAKQPRYPV-WQAKP-----EPQGVVL-EGA--- 214
 QY 371 PDDPSPTVTSVARKLRADINIVIGALGDSLTAGNGAGSTPGNVLDVLTQYRGLSWGVG 430
 DB 215 -----SSNPISTPOGNYIFLPGQ----- 232
 QY 431 DENIGVTTLANILRENPSPSKGPGVGTGKETSNAFLNQAAGRAEDLPVQARRLVLD 490
 DB 233 -----QERGLVLEAMVG-----YRATTIV 253
 QY 491 MKQDTRHNPEDWKITLFTGNDLGD-----FCNDLVHYSQONTNIGKALDILHA 543
 DB 254 FVSDTVLALSKVSGKELVMTAGKKQGAKEPSEILMTDGLVMTGRGTDDSG-TLOQLHI 312
 QY 544 EVPRAPV-----NLVTVLEIVNLRELVOEKVVCPRMILSLDPCVLKPEDNSTELATL 597
 DB 313 SPEKSYTLGDAGSGVFS-----NFFYESEITNRLYLFTDRPLRAGRDVVKV-- 364
 QY 598 IEFNKKFOEKTHQ-----LIESGRVDTREDFTVVQPFENVDMPTKSEGL--LPNSSFAP 651
 DB 365 --IGREHDPDLHSSPIVSAPAKSLVDANGSLLOTVAVTLDARRGGGSRFLPENAVAGG 422
 QY 652 DCFHSSKSHRASALMNMLEP--VGOKTRHKTE-----NKINITCPN----- 695
 DB 423 YELRLAARNQVSSSFFVANVYIKHFEIGLALAKKEKFGTGAVSGKQLDLYPDEPYNKA 482
 QY 696 QVQPFLETVNNSMGCHTWMPCRRAPASALHPSVHALRPADIQVVALDGLSLTAGCIG 755
 DB 483 RVQSLSLAQQLSMGN-----DLRYAGRPPVS-----LSSSEVSPASG 521
 QY 756 SKPDDLPDVTQYRGLSVSAGDGLSNVTLPLILREFNNTLGVAVGT-----GDANDT 811
 DB 522 HVALNLPRAAKPSRYLLTVLSASDGAAYRTTKELILE--RGLAHYELSLTAQVSNSES 579
 QY 812 NAFINQAVPQAK-----AEDLMSQVOTLMQKKMDHNVN-----HEDM 850
 DB 580 VFRFYALLESSKQVPVTEWMLRLDRTSHGELSPSGK--SETVNFAKPNVNLTLRDKG 638
 QY 851 KVTIVL---IG-----SDLCVCTDSNLYSAANFVDHLRNALDVLHREPRVNLVLD 901
 DB 639 LILAGLSHAIVSGKSTAHGTVDIVAAKTLVQPG-----ETAKMLITFPE 683
 QY 902 FLNPTIRQVFLGNPKCPVQQAIVLCNVC--LTLR--ENSGEILARLEAFSRAVSSMRE 957

DB 684 PIDEALL-----TLERDRVEQOSSLSHPRAMWTLQRIANDTOYEARVPV----- 726
 QY 958 LVSGRDYDIOEDFSVYL-----QPFQNIQCPVLADGLPDSFFAPADCIHNOKHSQLAR 1013
 DB 727 ---SNSFAPNITSVLTRNGQVSFOVAGIKVAVPOL-----DIRVTKDKHYPGE 775
 QY 1014 ALMTNMLEPISGKTEETDLR-----AEMPTCTQNEPPLRTPRNSNYT----- 1057
 DB 776 LVNVELTSLKGRFVASQLVGVVDEMIYALQELIANICKFFPYLGRNNVRRSSLSFI 835
 QY 1058 -----YPIKPAIENWGSDFLCTEWKASNVPTSVHOL-PPADIKV-VAALGDSLTITA 1107
 DB 836 SYDQALSSSEVADG-----ATNRSEBRVYKMLEPRRREEVDTAAMWPSLTLD 881
 QY 1108 VGARPNNSDLP---TSWRGLSWISIGDG-----NLEHTTILPNILKKNPY 1151
 DB 882 KQGAAYFTFLMPSLTFWRITARGMGDGLVGGRAVLRSEKILYMKMSPTVY----- 935
 QY 1152 ILGFSTWEGTAGLVNAAGARAPDPAQMDLVERMKSPDINLEKDWKLVTLFTGVN 1211
 DB 936 -----RVGDKRPAAGLFISSQDNBP-----VALYTKFAG-- 964
 QY 1212 DLCHYCENPEAHATEVY--QHIOALDILSEELP-----AFV-NVEVME 1255
 DB 965 --AEMRQTLTLHKGANVYSLTQNIQOS-GLLSAELQONGQODSISTKLSFVDSNMFVEQ 1021
 QY 1256 LASIYQGGKCMALAAQNNCTCLRHQSLSLEQELKNV-----W-----NLQHGIS 1303
 DB 1022 QKNVMLGGGNALMLPBOASNIRL--QSESTPQELFRNLDALVDEPFGVINTGSRLL 1078
 QY 1304 SFSSY-W-----HOYQREDAVVVQPFQNTLPLNERG-DTDLTFPSEDCFHSDGHA 1356
 DB 1079 PLSIAPRSIADHOSAAANDIRQMIQ--DNRLKMLQAGARFTWNGED-----GNG 1128
 QY 1357 EMALAMNNMLEPVGRKTTNNFTSHRAKLCSPSPSYLYTLRNSLLPDQAEAD-- 1413
 DB 1129 DAFILTAW-----AMVADWQASQALIGVQOPR--YQGHMIDS--YAEQADNMPLHL 1174
 QY 1414 --EVLVYA---VPVAGVGLVVGIGITVWRCRGGRRED 1448
 DB 1175 RALVLAWAQEMNLPCKT--LNGLEDAI--ARRGTIED 1209

RESULT 7
 SNE2 HUMAN STANDARD; PRT; 6885 AA.
 AC Q8XKH0; Q8NIS3; Q8NF49; Q8TER7; Q8MWM3; Q8MWM4; Q8MWM5; Q8WXH1;
 AC Q8WU50; Q8UFO4; Q8Y2L4; Q8Y4R1;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Neepirin 2 (Nuclear envelope spectrin repeat protein 2) (Syne-2)
 DE (Synaptic nuclear envelope protein 2) (Nucleus and actin connecting
 element protein) (NUANCE protein).
 GN SYNE2 OR NUA OR KIAA1011.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 ON NCBI Taxid=9606;
 RX [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 9), FUNCTION, CHARACTERIZATION, AND
 RP INTERACTION WITH F-ACTIN.
 RX MEDLINE=22113122; Pubmed=22118075;
 RA Zhen Y.-Y., Libotte T., Muncie M., Noegel A.A., Korenbaum E.;
 RT "NUANCE, a giant protein connecting the nucleus and actin
 RT cyoskeleton";
 RL J. Cell Sci. 115:3207-3222 (2002).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 4; 5 AND 7).
 RX MEDLINE=21652858; Pubmed=11792814;
 RA Zhang O., Skepper J.N., Yang F., Davies J.D., Hegyi L., Roberts R.G.,
 RA Weissberg P.L., Ellis J.A., Shanahan C.M.;

RT "Nesprins: a novel family of spectrin-repeat-containing proteins that
RT localize to the nuclear membrane in multiple tissues.";
RL J. Cell Sci. 114:4485-4498(2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22286983; PubMed=12408964;
RA Zhang Q., Ragnauth C., Greener M.J., Shanahan C.M., Roberts R.G.,
RT "The nesprins are giant actin-binding proteins, orthologues to
RT Drosophila melanogaster muscle MSP-300.";
RL Genomics 80:473-481(2002).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 6).
RC TISSUE=Testis;
RA Foustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX PubMed=12508121;
RA Heilig R., Eckenberg R., Petit J.-L., Fonknechten N., Da Silva C.,
RA Cattelico L., Levy M., Barbe V., De Berardinis V., Urea-Vidal A.,
RA Pelletier E., Vico V., Anthonard V., Rowen L., Madan A., Qin S.,
RA Sun H., Du H., Pepin K., Artiguenave F., Robert C., Cruaud C.,
RA Bruns T., Jallou O., Friedlander L., Samson G., Broctier P.,
RA Cure S., Segreus B., Aniere F., Samain S., Crespeau H., Abbaï N.,
RA Alach N., Boscus D., Dickhoff R., Dors M., Dubois I., Friedman S.,
RA Guyonoux M., James R., Madan A., Mairey-Estrada B., Mangenot S.,
RA Martins N., Menard M., Oztas S., Ratcliffe A., Shaffer T., Trask B.,
RA Vacherie B., Bellemere C., Belser C., Bernard-Gonet M.,
RA Bartol-Mavel D., Boutard M., Briez-Silla S., Comette S.,
RA Dufosse-Laurent V., Ferron C., Lechaplais C., Louesse S., Musielat D.,
RA Magdelienat G., Pateau E., Petit E., Silvain-Trukhietz P., Trybou A.,
RA Vega-Czarny N., Bataille E., Bluet E., Bordelais I., Dubois M.,
RA Dumont C., Guerin T., Haffray S., Hamadi R., Munga J., Pallouin V.,
RA Robert D., Wundler E., Gaugier G., Roy A., Saine-Marche L.,
RA Verdier J., Verdier-Discala C., Hillier L., Fulton L., McPherson J.,
RA Matsuda F., Wilson R., Scarpelli C., Gyapay G., Winkler P., Saurin W.,
RT "The DNA sequence and analysis of human chromosome 14.";
RL Nature 421:601-607(2003).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 8).
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhac N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.D., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whaley J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Fiering M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S., Krzywinski M.I., Skalek U., Smallus D.E.,
RA "Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM 3), AND SEQUENCE OF 1-956 AND 5133-6885
RP FROM N.A.
RC TISSUE=Spleen, and Tongue;
RA Jikuya H., Takano Y., Nomura N., Kikuno R., Nagase T., Ohara O.,
RA Ninomiya K., Wagauma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,

RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanenori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isozaki T.;
RT "NED0 human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE OF 5709-6885 FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=99246063; PubMed=10231032;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:63-70(1999).
RN [9]
RP REVISIONS.
RC TISSUE=Brain;
RX MEDLINE=22158633; PubMed=12168954;
RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
RT "Construction of expression-ready cDNA clones for KIAA genes: manual
RT curation of 330 KIAA cDNA clones.";
RL DNA Res. 9:99-106(2002).
RN [10]
RP SEQUENCE OF 5754-6885 FROM N.A.
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glaesli S.,
RA Ansoerg W., Boecker M., Bloeker H., Bautebachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
RA Mewes H.-W., Oestermeier B., Obermaier B., Poustka A.,
RA Mambrot R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
RN [11]
RP -1- FUNCTION: Involved in the maintenance of nuclear organization and
CC structural integrity. Probably anchoring protein which tethers the
CC nucleus to the cytoskeleton. Connects nuclei to the cytoskeleton
CC by interacting with the nuclear envelope and with F-actin in the
CC cytoplasm.
CC -1- SUBUNIT: Interacts with F-actin via its N-terminal domain.
CC -1- SUBCELLULAR LOCATION: Type IV membrane protein (potential). The
CC largest part of the protein is cytoplasmic, while its C-terminal
CC part is associated with the nuclear envelope, most probably the
CC outer nuclear membrane. Remains associated with the nuclear
CC envelope during its breakdown in mitotic cells.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=9;
CC Name=1;
CC IsoId=Q8WXH0-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q8WXH0-2; Sequence=VSP_007164, VSP_007166;
CC Note=No experimental confirmation available;
CC Name=3;
CC IsoId=Q8WXH0-3; Sequence=VSP_007155;
CC Note=Produced by exon skipping that results in a frameshift. No
CC experimental confirmation available;
CC Name=4; Synonym=beta;
CC Name=5; Synonym=alpha;
CC IsoId=Q8WXH0-4; Sequence=VSP_007156;
CC Name=6;
CC IsoId=Q8WXH0-5; Sequence=VSP_007157, VSP_007164, VSP_007165;
CC Name=7;
CC IsoId=Q8WXH0-6; Sequence=VSP_007158, VSP_007165, VSP_007166;
CC Note=No experimental confirmation available;
CC Name=8;
CC IsoId=Q8WXH0-7; Sequence=VSP_007154, VSP_007163;
CC Name=9;
CC IsoId=Q8WXH0-8; Sequence=VSP_007161, VSP_007162;
CC Note=No experimental confirmation available;
CC Name=9; Synonym=NUNACB-N-33;
CC IsoId=Q8WXH0-9; Sequence=VSP_007159, VSP_007160;
CC -1- TISSUE SPECIFICITY: Widely expressed, with higher level in kidney,
CC adult and fetal liver, stomach and placenta. Weakly expressed in
CC skeletal muscle and brain. Isoform 5 is highly expressed in

CC pancreas, skeletal muscle and heart.
 CC -1- DOMAIN: The Klarsicht domain mediates the nuclear envelope
 targeting.
 CC -1- SIMILARITY: Belongs to the Neprin family.
 CC -1- SIMILARITY: Contains 1 actin-binding domain.
 CC -1- SIMILARITY: Contains 2 calponin-homology (CH) domains.
 CC -1- SIMILARITY: Contains 1 Klarsicht domain.
 CC -1- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
 CC -1- SIMILARITY: Contains 9 spectrin repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; AF435010; AAL3547.1; -
 DR EMBL; AF435011; AAL3548.1; -
 DR EMBL; AY061757; AAL33800.1; -
 DR EMBL; AY061758; AAL33801.1; -
 DR EMBL; AY061759; AAL33802.1; -
 DR EMBL; AF495911; AAN60443.1; -
 DR EMBL; AL117404; CAB55905.1; -
 DR EMBL; AL162832; -; NOT_ANNOTATED_CDS.
 DR EMBL; AL155094; -; NOT_ANNOTATED_CDS.
 Query Match 1.64; Score 126.5; DB 1; Length 6885;
 Best Local Similarity 17.84; Pred. No. 67;
 Matches 219; Conservative 174; Mismatches 408; Indels 431; Gaps 56;
 QY 342 FSYRNSVYLRLQKPOKLEVRGAEIRCPDKPSDVP-----TSYRLKPADINVYGA 396
 DB 5598 FLVCCCKEMKIQLEKIEBALRV-----DVANSPELLLEQOKTYKMLEAVSINOT 5646
 QY 397 LGDSLFGANGAGSTPGVNLV-----LTOYRGAS--WS---VGDENIGVTTLAN 442
 DB 5647 IADSYV-----TOSLOLDTTELENRPETTESKLTDRKQNAVGVORQKQDVGLVR 5700
 QY 443 ILREFNPSLKGFSVGTGKETSPPNAFLNQAAGRAEDLPVQARLYDLMKQDRIHF--- 499
 DB 5701 QMODFTTSVENLFF--RPLTDSHLSAVKQGFSL-YOTRSLIHELKN-KEIHQR 5754
 QY 500 -----QEDMK-----ITLFGNDLQDF 518
 DB 5755 RTTCALTLAEGEKLLITTDLTKEVSGRISQLODSWKMEPOLAEMIKOFSTVETWQ 5814
 QY 519 C-----NDL-----VHSPONTDNGKALDILHA 543
 DB 5815 CEKKIKELKRLQVLAQSEDPRLBELHEDLHNEKELIKELBOSLASWTQNL-KELOTKA 5873
 QY 544 EVRAFAVNLVTLEIVNLR-----LYQEKVYCPRMILRSICPVLFKFDNSTELATLIE 599
 DB 5874 DLRRH-----VLVDVWVLKQEIHLHQWEDLCIRVAIRK-----QEIREDLNTWV 5921
 QY 600 FNKKFQF-----KTHOLIESGRVTRDFYVVOFFEN-VDMKTS 640
 DB 5922 FNEKNKELCAMLVOMENKVLQADISIEEMIKLOKQCMEEINL-----FSENNKQLQKMG 5977
 QY 641 EGPLDMSFFAPDCHFSKSHSRA-----SALMMNMLEBVGOKTRHKPENKINI 691
 DB 5978 DQI-----IKASNKSPRAEITDKLNKINDRMQHLFDVTGSRVKKLTK----- 6018
 QY 692 TCENOVOPFLRTYKNSWQHGQTLPCRDRAPSALHPTSVALRPADQVVAALGDSITAG 751
 DB 6019 -----ETPAFIQDQDKMNSNRTWLA-----RIESEL-----KRV-----VYDVCDDEIQ 6060
 QY 752 NGIGSKPDDLDTVTTQYRGSLSYAGDGSLENTTLLNLIIEENRNLTGYAVGTGANDT 811
 DB 6061 KRLAEQODLQDRIE-----QHSAG-----VESVFNIDVL-----LHDS 6094
 QY 812 NATLQNA-----VPGAKAEDLMQVQV--LMQKKKDHRVNFHEDMK 851

DB 6095 DACANETECDSIQOTTRSLDRWRNICAMSEMRMKIETETRLMQFLDDY--SREPDW- 6151
 QY 852 VITVLIGGSLDLCYCTDSN--LYSAN-----FVDHNRNL----- 885
 DB 6152 -----LKSARTAACNNSSEVLTSKBEKREAFQROHERLTQLELQKRYRLARE 6206
 QY 886 -----DVLHREVPRLVNLVFLNPTIRQVFLGPNDKCPVQAS 925
 DB 6207 NRTDTSRLKQWHEGQNRWINDQKRVAVLRLRFTN-----QREPEGRRE----- 6255
 QY 926 VLNCVLTLENSQELARLEAFSRA-YRSMRELVG-----SGRYTOEDPSVVLQ 976
 DB 6256 ---SILVWLTEMLOQLTNVEHFESDADDMRQLNGFOEITLNTKIDQLIVFGEOLIQ 6312
 QY 977 FPNQIQLPVADLPDTPSPAPDCIHPNOKFHSQ-----ARLMMNMLEPLG 1024
 DB 6313 KSEPLPAVLIEDELELHRYCOEVPGRVSRFHRRLTSCTPGLDEKASENETDMEDPRE 6372
 QY 1025 SKTETDLRAEMPICTQNEPEPLRTPRNSNYTPYPIKPALENMGS-----DFLCTEWA 1078
 DB 6373 IOTDSMRKGE-----SEEP--SSPQ--SLCHLYAPGHERSGCTPVSVDSI PLEMDH 6421
 QY 1079 SNSV-PTSVHQLRPADIKVVAALGDSLTGA-----VGAR-----P 1112
 DB 6422 TGDVGGSSSHBEDDEBPYVSALSGKSIDGSHMHPDPSCEPHNYKQMGDNVPEVP 6481
 QY 1113 NNSSDLEPTSWRGISMSIGDGNLEHTTLLNLIKKRNPPYLGGSTSTBSTAGLNAALG 1172
 DB 6482 ASSTPYKPPYKLLPBGTDGKEG--PRVLN-----GNPQOEDGGLADITEQSG 6530
 QY 1173 ARARDMPAQMDELVERMKNSPDI-NLEKQWVLTFLGVN-----DLCHYCEP-----BAH 1223
 DB 6531 AFRMENIQOELHANKIKIKONIQOLNSDISAITTKTEALEMLKAKPPSDIQEIE 6590
 QY 1224 LATEVYOHIOQALD-----ILSEELPRAFNVEVVELASLYOGGKCA 1268
 DB 6591 LRVKRLQELIKADTYKALVSVSVNSKEPLOTSESSESTELQRLQSLWME----- 6643
 QY 1269 MLAQNNCTCLRHS--QSSLEKQELKKNVNL 1298
 DB 6644 --AAQGVDSWGRGLRQSLMQCODFHQLSONL 6673
 RESULT 8
 ID ATCT_YARLI STANDARD; PRT; 928 AA.
 AC 043108;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Calcium-transporting ATPase 1 (EC 3.6.3.8) (P-type calcium ATPase).
 GN Yarrowia OR SCA1.
 OS Yarrowia lipolytica (Candida lipolytica).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Dipodascaceae; Yarrowia.
 OK NCBI_TaxID=4952;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98121320; PubMed=9461422;
 RX Park C.S., Kim J.-Y., Crispino C., Chang C.C., Ryu D.D.Y.;
 RT "Molecular cloning of YLMR1, a S. cerevisiae pMR1 homologue encoding
 a novel P-type secretory pathway Ca2+ ATPase, in the yeast Yarrowia
 lipolytica.";
 RT Gene 206:107-116 (1998).
 RL -1- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE HYDROLYSIS
 OF ATP COUPLED WITH THE TRANSPORT OF CALCIUM. HAS A ROLE IN THE
 SECRETORY PATHWAY.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+) (Cis) = ADP + phosphate +
 Ca(2+) (trans).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Golgi.
 CC -1- SIMILARITY: Belongs to the cation transport ATPase family (P-type
 ATPases).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U75447; AAC03419.1; -.
 DR HSP: P04191; 1EUL.
 DR InterPro: IPR006413; ATPase-11A2_Ca.
 DR InterPro: IPR001757; ATPase_E1-E2.
 DR InterPro: IPR006068; Cation_ATPase_C.
 DR InterPro: IPR004014; Cation_ATPase_N.
 DR InterPro: IPR000695; H_ATPase.
 DR InterPro: IPR005834; Hydrolyase.
 DR Pfam: PF00689; Cation_ATPase_C; 1.
 DR Pfam: PF00690; Cation_ATPase_N; 1.
 DR Pfam: PF00122; E1-E2_ATPase; 1.
 DR Pfam: PF00702; Hydrolyase; 1.
 DR PRINTS: PR00119; CATATPASE.
 DR PRINTS: PR00120; HATPASE.
 DR TIGR: TIGR01522; ATPase-11A2_Ca; 1.
 DR TIGR: TIGR01494; ATPase_P-type; 4.
 DR PROSITE: PS00154; ATPase_E1_E2; 1.
 KM Hydrolyase; Calcium transport; Transmembrane; Phosphorylation;
 KM Magnesium; ATP-binding; Golgi stack.
 FT TRANSMEM 81 101 POTENTIAL.
 FT TRANSMEM 105 125 POTENTIAL.
 FT TRANSMEM 271 291 POTENTIAL.
 FT TRANSMEM 306 326 POTENTIAL.
 FT TRANSMEM 718 738 POTENTIAL.
 FT TRANSMEM 750 770 POTENTIAL.
 FT TRANSMEM 799 819 POTENTIAL.
 FT TRANSMEM 860 880 POTENTIAL.
 FT TRANSMEM 887 907 POTENTIAL.
 FT MOD RES 353 353 PHOSPHORYLATION (BY SIMILARITY).
 SQ SEQUENCE 928 AA; 100050 MW; CBA53BDD075BEE7 CRG64;
 Query Match 1.6%; Score 125.5; DB 1; Length 928;
 Best Local Similarity 20.3%; Pred. No. 3;
 Matches 173; Conservative 120; Mismatches 317; Indels 243; Gaps 46;
 QY 220 VAEVSQYHGTWLSAPAPENCSEETRLAKVMQ-----WSQDEAN--SLASSRYSE 272
 DB 12 LSETNRSNHT--PSAQYQMTVEET-CSKLTNPETGLTSSQEMHRRDHIGSHFEQA 67
 QY 273 QESFTVFPFFYETTPSLHSEDPRLQ--DSTTLAWHLNRMWEPAGEKDEPLSY---- 325
 DB 68 EEDDSLI-KKFFQ-----FSENPILLILLIGAANVSFFM-----GNHDDAISTTLAI 113
 QY 326 ---KHGRPKCSQSPSLFSYRNSNYLRLQK--PODKLEVERG--AEIRCPDQPSD 377
 DB 114 LIATVVG-----FVQERYSEKSLKALNKLVPEAHLLRAGNSQVLASTLVPGD 162
 QY 378 TVPTSHRLKPADINIVIGALGDSLTAGNGSGTPGANVLTLYRGLSMSVGDEINIG-- 435
 DB 163 LVEFSVGDRIPACRIVKAVHLSIDSNLGETTPVTKD-----TNPVSTPPIGLA 214
 QY 436 ---TVTTLANILREPFSLKGFVGTGKETSPPAFLNQAVAGRAEDLPQARLVLDWK 492
 DB 215 DRNTAVMGVLVDNGGT--GIYVGSGSHAPAVYDM-VSEISTKPTPLQ--SMIDLK 270
 QY 493 NDRTHIFQEDWKI--TTLFGNDLDCDFCNLDVHVSPOFTDNIGALDILHAVERPAFV 550
 DB 271 DLSIVFGVGVICLIGMFGGRDML-----EMPT--IGVSLAV--AAIRGGLP 314
 QY 551 NUTVLEIYNLRLEQEK-----KYVC--PRAILSLCC-----583
 DB 315 IIVTVLALGVLSRQKAIVRKLPSVETLGSVNVVICSDDKTLTLTRNHSCTTCWTVNG 374
 QY 584 -----VLKFDNSTELATLIER-----NKKFEKTHQILIES- 614

DB 375 DLANAVTLKPGSHTEADPKVAALK--NSVSLANMLKVNLCNNSKFNREAGHLYGNA 431
 QY 615 -----GRYTREDFTVYVQ--PPFENVDMKTSBGLPDS-----FFA 650
 DB 432 TDIALIEVDLYFLEDTRETRKVAEPFSSSRKMLTSTTTGDSSTPMISVKGAGEVIA 491
 QY 651 POCFHSKSHSRA-----ASALNMMLPEPQCKTRFKFEKINITCPNOV 697
 DB 492 PFCETVCKDOKTAPPNDDMRKVTETIASBMSDGLRIIAFAVKQKYE-----GSEE 545
 QY 698 QPELRTYKNSMOGHGTWLPQRDRAPSAHPPTSVAHLPADIQVVAALGDS---LTAGNG 753
 DB 546 APEGLVFAGLM--GLYDPRDPVRA-----IRRLTTGSVRVMTTGSAAATLISGR 597
 QY 754 IGSKPDDLVDVTTQYVGLSYSGAGDSLEN-----VTLF---NIRENRNLTY 801
 DB 598 IGWPL--MPGTOSVSGSKLATWSQALDECTOTASIPARTSPEDKIKVKGFOR--GD 653
 QY 802 AVG-TGD-ANDTVAFLNQAVPGAEDL--MSQVOTLQKMKDHRVNFHEDMKVTTL 856
 DB 654 VVAMTGDVND-----APALKLADIGANGGGTIVAKAADMILTDDEATTLSAI 705
 QY 857 IGSSDLCDCYCTDSNLYSAANFVD-HLRNALDVLHREPVRLVNLVDFLNP--TIMQVFL 913
 DB 706 EESKGIFFN-----NINFTFQSTSMALSLYAAVATIMGLEPPLNMQITLNIIM 757
 QY 914 GNPDKCPVQOASY 926
 DB 758 DGP---PAQSLAV 767

RESULT 9
 MN4 YEAST
 ID MN4 YEAST STANDARD; PRT; 1178 AA.
 AC P36044; P36043; P89095;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE MN4 protein.
 GN MN4 OR YKL200C/YKL201C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxId=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c;
 RX MEDLINE=97175967; PubMed=9023541;
 RA Odani T., Shima Y.-I., Tanaka A., Jigami Y.
 RT "Cloning and analysis of the MN4 gene required for phosphorylation
 of N-linked oligosaccharides in Saccharomyces cerevisiae";
 RL Glycobiology 6:805-810(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Maia e Silva A., Boseler P., Vilela C., Fernandes L., Soares H.,
 RA Guerreiro P., Rodrigues-Pousada C.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MAY FUNCTION AS A POSITIVE REGULATOR FOR
 CC MANNOSE-6-PHOSPHATE TRANSFERASE. IS REQUIRED TO MEDiate
 CC MANNOSE-6-PHOSPHATE TRANSFER IN BOTH THE CORE AND OUTER CHAIN
 CC PORTIONS OF N-LINKED OLIGOSACCHARIDES.
 CC -1- SUBCELLULAR LOCATION: TYPE II membrane protein (Probable).
 CC -1- SIMILARITY: TO YEAST YOR061M.
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
 CC FRAMESHIFTS, ONE OF WHICH PRODUCES TWO SEPARATE ORFS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@slb.ch).

CC -----

DR D83006; BA11676.1; -

DR EMBL; Z28201; CA82046.1; -

DR EMBL; Z28200; CA82044.1; -

DR PIR; S78475; S78475.

DR SGD; S0001684; MNN4.

DR GO; GO:0006487; P:N-linked glycosylation; IMP.

DR GO; GO:0006493; P:O-linked glycosylation; IMP.

DR GO; GO:0006950; P:response to stress; IDA.

KW Transmembrane; Signal-anchor; Repeat.

FT TRANSMEM 1 27 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 28 48 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)

FT 49 1178 (POTENTIAL).

FT DOMAIN 1032 1174 ARG/GLU/LYS-RICH (HIGHLY CHARGED).

FT DOMAIN 1042 1174 17 X 8 AA TANDEM REPEATS OF K-K-K-E-E-E-E.

FT REPEAT 1042 1049 1.

FT REPEAT 1050 1057 2.

FT REPEAT 1058 1065 3.

FT REPEAT 1066 1073 4.

FT REPEAT 1074 1081 5.

FT REPEAT 1082 1089 6.

FT REPEAT 1090 1097 7. (APPROXIMATE).

FT REPEAT 1098 1105 8.

FT REPEAT 1106 1113 9. (APPROXIMATE).

FT REPEAT 1114 1121 10. (APPROXIMATE).

FT REPEAT 1122 1129 11. (APPROXIMATE).

FT REPEAT 1130 1137 12.

FT REPEAT 1138 1144 13. (APPROXIMATE).

FT REPEAT 1145 1152 14. (APPROXIMATE).

FT REPEAT 1153 1160 15. (APPROXIMATE).

FT REPEAT 1161 1168 16. (APPROXIMATE).

FT REPEAT 1169 1174 17. (INCOMPLETE).

FT REPEAT 37 40 POLY-ILE.

FT DOMAIN 1042 1045 POLY-LYS.

FT DOMAIN 1046 1049 POLY-LYS.

FT DOMAIN 1050 1053 POLY-LYS.

FT DOMAIN 1054 1057 POLY-LYS.

FT DOMAIN 1058 1061 POLY-LYS.

FT DOMAIN 1062 1065 POLY-LYS.

FT DOMAIN 1066 1069 POLY-LYS.

FT DOMAIN 1070 1073 POLY-LYS.

FT DOMAIN 1074 1077 POLY-LYS.

FT DOMAIN 1078 1081 POLY-LYS.

FT DOMAIN 1082 1085 POLY-LYS.

FT DOMAIN 1086 1089 POLY-LYS.

FT DOMAIN 1094 1097 POLY-LYS.

FT DOMAIN 1098 1101 POLY-LYS.

FT DOMAIN 1102 1105 POLY-LYS.

FT DOMAIN 1134 1137 POLY-LYS.

FT DOMAIN 1157 1160 POLY-LYS.

FT DOMAIN 1165 1168 POLY-LYS.

SQ SEQUENCE 1178 AA; 139380 MW; BC05DAE0AEFCB282 CRC64;

Query Match 1.6%; Score 125; DB 1; Length 1178;

Best Local Similarity 18.0%; Pred. No. 4.8;

Matches 198; Conservative 134; Mismatches 317; Indels 448; Gaps 54;

QY 256 SYOEAANSLLASRYSEQSEFTVVFQFFYE-----TTPSLHSE----- 294

DB 69 SDDDSWYSILTSSFKNDK--IOPAKTLYENLKFGNPKWVNEYTLQNDLLSYKMPRK 125

QY 295 -----DPLQDSTTL-----AWILW----- 309

DB 126 GSKLESVDLKFYFDPLRTWSVLNLQNNADADQPKLPSWSYDWTTPHFLNKLISDK 185

QY 310 -----NMMPEAGKD--EPLSVKGRPMKCPSESEYLLSYNSNTLTRI-----QKPOD 358

DB 186 TVLPNCNLFQSAFQKESLEAIETELGEP-----LFLYERPKYAQKLMYKAAHNP 235

QY 359 KLEVRREGAEIRRC-----PD-----KQPSDVPTSVHRLKPADINVIGAL 397

DB 236 RIKDSKELKKGSKLFTPPDHGSPKGLRFNTQFOIKELYDKRPEVYQOAR--NYI----- 290

QY 398 GDSLTGANGAGSTPGVNLVLTQYRLSWSVGDEIGIVTTL-----ANILREF 447

DB 291 -----LTTQSHPLSISTESDSNVYQVLPQTEKSKNLVQSGLOEY 331

QY 448 -NPSLKFSVGCTGKETSPPAFLNQAVAGRAEDLPQARRLVDLMKDPRIFHQEDMKII 506

DB 332 INDNIN-----STTKRKK-----NKQDVEFNHRLQEPFNNOVN-----SLY 370

QY 507 TLEFGNDLCDFCNLDVHYSPOFTDNICKALDILHAEPRAVNLVTVEIVNLELYQ 566

DB 371 KLEIETDKEFTPDKDLVLPSPDFKFDASKIE-----ELBE 407

QY 567 EKVYVPRM-----ILRSLCPV-----LKF-----DNDSTELAT 596

DB 408 QKKLYPDKSAHNENYLNLSKSVKTSPLQRFFYAGAVYQYKMGFHRDRFFNVIT 467

QY 597 LIEFNKKFOEKTQHLIESGRYDREDFTVVOFFENYDMPKTSGLPDNSFFAPDCFH 656

DB 468 LINDKQEYQARLNSMR-----TFQKT-----KANGI-----I 496

QY 657 SSRSHRAASALMNNLPEVGOCTTHKKEKNKINITCPNOVQFLRTYNS-----MQ 709

DB 497 SWLSHGTLGYLYNGMAP-----WDNDFDLQMPKHLQLLSQYFQSLILEDP 547

QY 710 GHGTWLPCKDRAPSAHPHSVHARPADIQVAAAGDSILTA---GNGISKPDLDPV 766

DB 548 GNGRYP-----LDVSDSLTVRINGGKNNDARFLDVT 581

QY 767 QYRGLYSAGGDSLENVTTLPNILAEFNRLTGYVGTGADNDTAPLNOAVPGAKAD 826

DB 582 ---GLYI-----DITGLASTAPSRD---YANSYEELOE 612

QY 827 LM-----SQVQTLQMKQKDDHRVNFHEMDKVTTLIGSDLCDYCT-DSN-----L 871

DB 613 HLDINNIPESNGETATLPKV-DDGLVM-----ATLITELADYITSDENKHKEY 663

QY 872 YSAANVDHRLKRLADVLHREVRVLYNLVDPLNPTIMROVFLGNPKCPVQASVLCNCV 931

DB 664 PTDIDKDLKKELE---ELPK-SKTIENKLP--KQRFYNEKXK-----LYNC- 707

QY 932 LTLRENSQELARLEARSRAVRSRRELVSGRYDQEDSFVVLQPFQONITQPLVLDG 991

DB 708 ---RNN-----HNSFEELSPLINTYFHGV--PAL---IP 734

QY 992 DTSFAPDCIHPNQKHSQALRALMTNMLEPLGSKTETLIDRAEMPTITCPTONEPFLRTP 1051

DB 735 HRRTY---CLHNEYHVDRYAFDAYKN-----TAYLDEFRFWDYDGLKIC 777

QY 1052 RNSNYTPYIKPALENNGSDPL---CTEWKA---SNSVPT--SVHQLRADIKV----- 1097

DB 778 SNINSWYPMNPISINSWNPILKEISSTKFSKLDNSKNKSEYFKNLSMDVDVLLKNIP 837

QY 1098 -AALGDSLTAVAGARPNNSSDLPSTSRGLSWSIGCGDNT-----THTLLPNILKFF- 1148

DB 838 KAFIEVFYTLV-----NSPV-TAYRQKELEIYQCNLTFIEKKLLQRLINVAPKLS 891

QY 1149 ---NPLYLGFSTIWE 1161

DB 892 SPAKDPFLGYEKAMWK 908

RESULT 10

YFAS ECOS7

ID YFAS ECOS7 STANDARD; PRT; 1534 AA.

AC 08XK35;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein yfas precursor.

GN YFAS OR Z3481 OR ECS3111.

OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia..
 OC NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / BDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=1206551;
 RA Perna N.T., Plunkett G., Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grodeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=1258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takem H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hatiori M., Shingawa H.;
 RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 CC -1- SIMILARITY: BELONGS TO THE UPF0192 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AE005454; AAG57357.1; -;
 DR EMBL; AP002560; BAB36534.1; -;
 DR PIR; G91017; G91017.
 DR KMW Hypothetical protein; Signal; Complete proteome.
 FT SIGNAL 1 38
 FT CHAIN 39 1534 HYPOTHETICAL PROTEIN YFAS.
 FT SEQUENCE 1534 AA; 169893 MW; E326FD052770F1A3 CRC64;
 Query Match 1.6%; Score 123.5; DB 1; Length 1534;
 Best Local Similarity 18.5%; Pred. No. 9.3;
 Matches 274; Conservative 179; Mismatches 554; Indels 475; Gaps 66;
 QY 161 OPDWKLIINFSSNASQCYLCPSAQNGLA-----AGVDLMLGVLDTLQOE 206
 DB 9 QHFMHL-SRFGSGLAIACLSLVGTGLNADSLPSSNVPAPAGTFFLLDSSFSSE 67
 QY 207 VPRAFYNLVDLSEVAESRQYH-----GTWLSPAPEPCNCEETTRLAKVVMQMSY-- 257
 DB 68 ESK-----VRLAEPGDYRKYQMEYGVQDVRLYRIPDEMAFLRQOKLHRIYVQPOYIG 122
 QY 258 -----QAWNSLASSRSESEFTVVPQPFYETTSLSHSDPRLODSTTLAWHLNN 310
 DB 123 DGLNLTLMWLMDDWYKSRVRVQRTFSSQGRQNVTOALPEL-----QLGN 167
 QY 311 RMMEPRAG-EKDEPLSYKRGPRMKCPQSEPSYLESYNSNVTLRLOKRODKLEVRBGAER 369
 DB 168 AAIKPERYVONNOFS-----PLKQYPLVQOFYPLMQAKPVEPQGVKL-BGA--- 214
 QY 370 CPDKDSDVTPVTSVHLKPADINVIGALGDSLTAAGAGAGSTPGNVADVLTYQGLSMSYG 429
 DB 215 -----SSNFISQPGNIIYIPLGQ----- 232
 QY 430 GDENIGTVTTLANILREPNPSLKGFSVGKETSPPAFLNQAVAGRAEDLPVQARLVLD 489
 DB 233 -----QEPGLYIVEAMVGG-----YRATTIV 252

QY 490 LMKNDTRHFOBDWKIITLFIGNDLCD-----FCNDLVHYSQNFNTDNIGKALDIH 542
 DB 253 VFSVDTVLALSKVSGKELVWTAGKKQGEAKPSEILMTDGLGWMRTGVTDDSG-TLQLQH 311
 QY 543 AEVPRAFV-----NLVTLLEIVNLRELQOEKKVYPPRILSLCCVLFKFDNSTELAT 596
 DB 312 ISPERSYILKDAEGGVFVSE-----NPFSEBIYVTRLYIFTRDLRYAGRVQVKKV-- 364
 QY 597 LIEFNKKFOEKTHQ-----LIESRGYDTRREFVTYVVOFFENVMPKTSEG---LPDMSFFA 650
 DB 365 ---MGEHFDPHLSSEFYSAAPKLSVLDANGSLQLQVDVTLTDARRNGQGSFRLPEAVAG 421
 QY 651 POCFFSSKSHSRASALNNMLER---VGOKTRHKFE-----NKINITYCPN----- 695
 DB 422 GVELRLAYNQYSSSFRVANYIKHPELGLALDKKEPFTGEAVSGKLQLLYPDGSPVK 481
 QY 696 -QVQPFRLTYKNMOSQHGTMWPCRDRAPSALHPTSVHALRPADIOVVAALGDSLTAANGCI 754
 DB 482 ARVQLSLRAQQLSMVGN-----DLRYAGREFVVS-----LEGSETVSDAS 520
 QY 755 GSKPDDLPVTTQYGLSYAGDGSLENTVTLPLNREPNLTGYAVGT---GDAND 810
 DB 521 GHVALNLPPADKPSRYLLTVSASDGAAYRVTTKEILIE-RGLAHYSUSTAAYNSNGE 578
 QY 811 TYAFLNQAVPGAK-----AEDLMSQVOTLMQKMDHRVNF-----HED 849
 DB 579 SVFFRYALLESSKQVPVYTEMRLBERTSHSGDLPSGCK-SFTVYNDKGNVNLTLRDKD 637
 QY 850 KWVITVL-----IGSSDLCDYCTDSNLYSANFVDHLRNALDVLREVPVLNVLV 900
 DB 638 GLILAGLSHVGKSGMSHTGVDIADVADKTLVQPG-----ETAKMLITFP 682
 QY 901 DPLNPTIMQVQLGNPDKCPVQOASVLCNCV--LTIR--ENGELARLAFSAVYSSNR 956
 DB 683 EPIDEALL-----TLERDRVEQSLSLSPANWMLTLQRINDYQEARVVP----- 726
 QY 957 ELVSGRYDTPQEDFSVL-----QPFQNIQLPVLADGLPDTSFAPDCTHPNOKFHSOLA 1012
 DB 727 ---SNSFAPNITFSVLTYRNQYISQNGIKYAVVQL-----DIRKDKTHYQPG 774
 QY 1013 RALMTNMLEPILGSKTETTLDR-----AEMPTICPTQNEBFLTPRNSNYT----- 1057
 DB 775 ELVNVETLSLKGKPPSAQLTVGVDEMITYALQPEIAFNIGKFYPLGRNNVRTSSLSLF 834
 QY 1058 -----YPIKPALEMGSDPLCTEMKASNVPSYHQL-RPADIKY-VAALGSLT 1106
 DB 835 ISVDQALSEEPVAPG-----ATNRSERRVWMLRPREEVDTAAMPSTLT 880
 QY 1107 AVGARPNNSDLP---TSWRGLSMISIGDG-----NLEHTTLPLILKKFNP 1150
 DB 881 DKQKAYFTFLMPDLSLTRITARGNGDGLVGQRAYLRSEKNLYMKMSMPTVY----- 935
 QY 1151 YLGFSTSTWEGTAGLVNAABGARADMPAQMADLVERKNSPDIINLEKDWKLVTLFIGV 1210
 DB 936 -----RVGDKPSAGLFIISQQDNB-----VALVTFAG- 964
 QY 1211 NDLCHYCEPBEALATEY---OHIOQALDIISEELR-----APV-ANVEVM 1254
 DB 965 ---AEWRQTLTLHGKANYISLAQNIQOS-GLISAEIQQNGOVODSISTKLSEFVDSNPVE 1020
 QY 1255 ELASLYOGGCGKAMLAQNNCTCLRHQSSELEKEKLVN-----W-----NLQHG 1302
 DB 1021 QOKNVNLGGDNALMLPEQASNIRL---OSSETPOEIPFNNDALVDEPMGVINIGSR 1077
 QY 1303 SSFSY-W-----HOYTQREDFAVVVOFPQNTLTPLNBERG-DTDLTFSEDECFHPSDRGH 1355
 DB 1078 IFLSLAMRSLADHQSAANDIRQMIQ---DNRLRLMQGLGPGARFTWMGED-----GN 1127
 QY 1356 AEMALALNNMLEPVGRKTSNNFTHSRAKLKCPSESEYLYTLNLSRLLPQAEAP-- 1413
 DB 1128 GDAFLTAW-----AMVADWQASQALGVTOQPE--YMOHMDS--YAOQADNMPPL 1173
 QY 1414 ---EVLVMA---VPAAGVGLVVGIIQTVVWRRCRGRGRED 1448

Db 2132 -TUL-----DEVT-IPNFVSHALLSLD---DEDQVNVNGFNALS-TLKKVD 2174
 Qy 913 LGNDKDC--PVQASVLC-----NCVLT-----RENSOELA 942
 Db 2175 KPTLEKLVKAKOSLALTGOCODVAFKLPRGNCVLPFLIGMVGNDDEEBS-ALA 2233
 Qy 943 RLEAFBRVAVRSKRELVGSGRYPTQEDFSVLPQPFQCNQLPYLADGLPDTSEFA----- 997
 Db 2234 IADVSKTPAANKPFVS---VITGPLIRVGERFSSDIKAIL-----FALNVLF 2281
 Qy 998 ----PDCIHPRQKHSQALRYALNTMMLPGLSKTETTLDRAMPITCTQNEP 1046
 Db 2282 IKIPMLRP---FIPOLQRTFVXSLD---ATNETTLRLAAKALGALIEHP 2327
 RESULT 12
 PDP2_ARATH STANDARD; PRT; 1039 AA.
 ID PDP2_ARATH
 AC 09M9W8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Phospholipase D p2 (EC 3.1.4.4) (AcPLD2) (Phospholipase D2 PHOX and
 PK containing domain) (Phospholipase D zeta 2) (PLDzeta2).
 GN PUDP2 OR AT3G05630 OR F18C1.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eurosidia II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxId=3702;
 RX MEDLINE=21016720; PubMed=11130713;
 RC STRAIN=cv. Columbia;
 RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseld M.,
 R Farnham B., Vailie G., Bloecker H., Perez-Alonso M., Obermayer B.,
 RA Delseny M., Boutry M., Griwall L.A., Mache R., Pilsdomelech P.,
 RA De Simone V., Cholsene N., Artiguenave F., Robert C., Broctier P.,
 RA Wincker P., Catolico L., Weisenbach J., Saurin W., Querrier F.,
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
 RA Reichelt J., Scharte M., Schoen O., Bargues M., Terol J., Clement J.,
 RA Navarro P., Collado C., Perez-Perez A., Ottensmeider B., Duchemin D.,
 RA de Haan M., Maarte A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
 RA Monfort A., Argirou A., Flores M., Ilguori R., Vitale D.,
 RA Manhaupt G., Haase D., Schoof H., Ruid S., Zaccaria P., Mewes H.-W.,
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
 RA Rooney T., Rizzo M., Walters A., Utecherback T., Fujii C.Y., Shea T.P.,
 RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
 RA Pal G., Miltcher J., Sellers P., Gill J.E., Feldblum T.V.,
 RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
 RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asanizu E.,
 RA Sasamoto S., Kimura T., Ideesawa K., Kawashima K., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.,
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:820-822(2000).
 CC -1- FUNCTION: Hydrolyzes glycerol-phospholipids at the terminal
 phosphatidylesteric bond. Phosphatidylcholine-selective.
 CC -1- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = choline + a
 phosphatidate.
 CC -1- ENZYME REGULATION: Calcium-independent and PIP2-dependent.
 CC -1- SIMILARITY: Belongs to the phospholipase D family. PXP-PLD
 subfamily.
 CC -1- SIMILARITY: Contains 1 phox homology (PX) domain.
 CC -1- SIMILARITY: Contains 1 PH domain.
 CC -1- SIMILARITY: Contains 2 PLD phosphodiesterase domains.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL, outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AC011620; AAP26134.1; -
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR001736; PLD.
 DR InterPro: IPR001683; PX.
 DR Pfam: PF00169; PH; 1.
 DR Pfam: PF00614; PLDC; 2.
 DR SMART: SM00233; PH; 1.
 DR SMART: SM00155; PLDC; 2.
 DR PROSITE: PS50003; PH DOMAIN; 1.
 DR PROSITE: PS50035; PLD; 2.
 DR PROSITE: PS50195; PX; FALSE NEG.
 KM Hydrolyase; lipid degradation, Repeat; Multigene family.
 FT DOMAIN 45 205
 FT DOMAIN 215 343
 FT DOMAIN 465 492
 FT DOMAIN 840 867
 FT ACT_SITE 470 470
 FT ACT_SITE 472 472
 FT ACT_SITE 477 477
 FT ACT_SITE 845 845
 FT ACT_SITE 847 847
 FT ACT_SITE 852 852
 SQ SEQUENCE 1039 AA; 117902 MW; A8433C237C3B7F73 CRC64;
 Query Match 1.6%; Score 123; DB 1; Length 1039;
 Best Local Similarity 18.7%; Pred. No. 5.4;
 Matches 157; Conservative 123; Mismatches 313; Indels 246; Gaps 37;
 412 GNVLDTLT-QYNGLSVSGDENIGVTTLANTLRFPNSLKGFSVGTKET----- 462
 273 GKLDLWPTDTGLQ---GTSSSQPLRLAQVKEHNLRFGFKVSGDRTVRLRTSS 328
 463 -----SPNAFLNQAAGRAEDLPVQARLYD-----LMKNDTR 496
 329 RYKVEWVAVDAGCGSPHRCFSPAPRLTSD-GSQAMPWDGHTAFALMAGLNATS 387
 497 IHFOEDWKITTLFIGNDLDCFDNDLVHVSPOFTDNICKALDILHAEPRAFNLVTVL 556
 388 EIFMTGWWL-----CPRL-YLKRPFEDHPSLRDLALLETAKGVKEVQI- 431
 557 EIVNRELVOEKKVYCPRMILSLCPCVAKFPDNLSTELATLLEFNKKPQEKTHQL----- 611
 432 -AKKNSLTSKRLQNIHNKVK---VLKRPDHLSSGIYLSNHEKIVYDVQVFCIG 485
 612 ---IESGRYTRDPTVVQPF---ENVMPKTSBGLPNSFPAPDCPFSSKSHR--- 663
 486 LDLCFRYVTAHKIGDCPRYIWPCKDYNNPESR---FNSWETMKDELDRKKYRMPW 542
 664 --AASALNNMLBPVGOKTRRKFNKINITCPNQVQPLRYUKNSMQHGTVLPCRDRA 721
 543 HDVHCALMPRCRDVARNHFGVQWNSHKRKAPREQITPLRMHNVLPB--YLGTRE-- 598
 722 PSHLHTSVHALRP--ADIVVAALGDLTANGSGSRDDLPVTVQYRGLSYAGGD 778
 599 ----IDITIAAKPPEEDPKPVLLAHDSFS---ASPPQELPILLPQETDAGFRGD 649
 779 -----GSLNVTTLPLNLRF-----NRNLGVAVGTGDANDT-- 811
 650 LKUDSGARODPPTSESLSDEAVNDMMWQIGKSCRCQIIRSVGQWAGSISQPDSTH 709
 812 -----NAFLNQAVGAKAED--LMSQVOTLMQKMDHRVNFHEDWKV 852
 710 RAVCSLIGNAHEHYIENQF---ISGLEKEDTILNRVLEALYRRI-----LKAHEENKC 761

QY 853 ITVL1-----GGSDLCDYCTDNTLSANFVDHRLNALDYLHREVPRLVNLVDFL 903
 Db 762 FRVIVIPBLRGCGGID-----DFGAT-----VRALMHOYTIRREGISIDNL 808
 QY 904 N-----PTMRQVFLCN-----PDKCPVOQASVLCNCVL 932
 Db 809 NALGPXTQDIISFYGLRSGYRLEFEDGPATISQIVHASKMIVDRIAVIGSSINIRSL 868
 QY 933 TLRENSOELARLE--AFSRAVRSRMRVLGSGRDTC-----EDPSVVLQPFQNIOLPV- 985
 Db 869 LGRDSIEIGVIVIEKFEVESNMGKMKWAGKFSYLSGMSHGLHAGEIQKIEDPIK 928
 QY 986 -----LADGPDTSFPAP--DCIHPNOKFHSQALRALMTWML-----EPYGSKTETLIDL 1032
 Db 929 DATYKOLMATAKNTDITVQVFSCL-PNE--HIRSRPALRHNMALCKDLGHTTIDLGI 985
 QY 1033 RAEMPTCPQNEBFLTPPNSNTYVPIKPAIENWGSDFLCTEWMKSNVPTSVHQLRP 1091
 Db 986 APERLESCGSWSIELKTRGNLVCPLQ-----FMCDQ-----EDLRP 1024

RESULT 13
 GALK_SCHPO STANDARD; PRT; 713 AA.
 AC Q9HJ3;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE GAL10 bifunctional protein [includes: UDP-glucose 4-epimerase
 (EC 5.1.3.2) (Galactowaldenase); Aldose 1-epimerase (EC 5.1.3.3)
 (Mutarotase)].
 GN GAL10 OR SPBPB2B2.12C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC NCBI_Taxid=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; Pubmed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Scouras K., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moulé S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Shelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voickert G., Aert R., Roben J., Grymopiez B.,
 RA Welfens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moselel D., Hilbert H.,
 RA Beyer P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey P., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cornetti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -!- FUNCTION: Mutarotase converts alpha-aldose to the beta-anomer. It
 CC is active on D-glucose, L-arabinose, D-xylose, D-galactose.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -!- COFACTOR: NAD.

CC -!- PATHWAY: Galactose metabolism; third step.
 CC -!- PATHWAY: Hexose metabolism.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE SUGAR
 CC EPIMERASE FAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE ALDOSE
 CC EPIMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: AL512522; CAC21414.1; -.
 CC HSP, P08147; 1XEL.
 CC GeneDB: Spombe; SPBPB2B2.12C; -.
 CC InterPro: IPR001823; Ald1 epimerase.
 CC InterPro: IPR001509; Epimerase_Dh.
 CC InterPro: IPR005886; GalE.
 CC InterPro: IPR000205; NAD_binding.
 CC Pfam: PF01263; Aldose_epim; 1.
 CC Pfam: PF01370; Epimerase; 1.
 CC TIGRPFAMs: TIGR01179; galE; 1.
 CC PROSITE: PS00545; ALDOSE_1_EPIMERASE; 1.
 CC KEGG: Multi-functional enzyme; Isomerase; NAD; Galactose metabolism.
 CC FT DOMAIN 1 350
 CC FT DOMAIN 351 713
 CC FT NP_BIND 7 38
 CC FT ACT_SITE 532 532 MUTAROTASE (POTENTIAL).
 CC SEQUENCE 713 AA; 80666 MW; 51689DA0843A8556 CRC64;

Query Match 1.6%; Score 122.5; DB 1; Length 713;
 Best Local Similarity 19.8%; Pred. No. 3.2; Indels 231; Gaps 38;
 Matches 138; Conservative 92; Mismatches 235;
 QY 370 CPDLPDSD-----TVPTSVHRLKPAD-----INVIGALDSDSLTAGAGSTPG 412
 Db 149 CPN-DPTNPGKTKVAIENIKDLHTSDNTWRGAILRYNPIGAKHSGGLGEPLG-IPN 206
 QY 413 NVLDVLTQY---RGLSMVQGD---NIGT-----VTTLA-----NIREFNP-- 450
 Db 207 NLLPFLAQLVAGREKLVFGDDYDSHDGTPRIDYIHVVDLAKGHIALNLYLNKINSSEG 266
 QY 451 -LKGFSVGTCKETSP---NAPLQAVAGRAEDLP--VQAR--LVLDKMDTTHIQ 500
 Db 267 MYREWMVGTKGSSVFDIYHAFCKEV---GDLPEYVAGRTGDLMLTASPNRANSE 321
 QY 501 EDKITTLLFGNDLDCDFCNDLVHVSPOFTDNIGKALDILHAEVPRAFNVLTVLEIYN 560
 Db 322 LKKK-----AEISITDACHDLMKWTEN-----PFGF-----QIDN 352
 QY 561 LR-ELYOEKKVYCPRMILSLCPVLKPDNSTELATLIEFNKFOEKTHQLI-----ESG 615
 Db 353 YKKKLTGTLGIMQYKRLHTIC-----FQDLVSIANYGALVQAVYKGRNLVNGFNDFS 407
 QY 616 RYTRREDFTVVOVPPFENVDMPKTSGLPDNSFFADCFHFSKSKSRASALMNNMLP 675
 Db 408 RYKLKEN-----PFF-----GATIGFAMRIANQOFEV 435
 QY 676 VQGTTRHKPE-NKINITCPNQ-----VQPELRYK--NSMQ-----GHGTWLPGR 718
 Db 436 DGLHLYTLCKENKNTLLHGNGNCFDQFGLGIARQYEDYNTLFLVDKGNNGF----- 491
 QY 719 DRAPSAHLPTSVVALRPADIQVVALGDSITAGNGISGKPDLPDYTTYRGLSYAGGD 778
 Db 492 ---PSDELITVKTITKNNLSL-----EKSVIPEYSKL----- 522
 QY 779 GSENVTTLPNLIREFNRNLTGYAV-----GTGDANDTNAFLN-QAVPGAKAED 826
 Db 523 ---NTAV-----NLTNHSYVNLASPNKTIIDGTIIKSTYNYLVKVNSETSPGTG 569

QY 827 LMSQVOTLMQKMDHVRNFHEHDKVITVLIGSSDLCDYCTDSNLYSAANFDHLRNALD 886
 DB 570 IYEMQNDITKPTLDIPDISD-----NCFIVDRASKSCFLDTRKTS-----LNKIVE 616
 QY 887 VLHREPRVLVNLVYDLNLPIMQVFLGN-PDKCPVQASVLCNCLTLRENSQELARLE 945
 DB 617 VHPSPVVKLV--VSTTEPAF--QLYTGDNDCICEQSRSGFC-----VETGRFT 662
 QY 946 AFSRAYSSRMELVSGSR--YDQEDFSVVLQPFQFN 980
 DB 663 NALNNEKSKQVILRKGEVYGARSKFSLVADLEEN 698
 RESULT 14
 SC71 SCHPO STANDARD; PRT; 1811 AA.
 ID 09UT02; 014168;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DE Protein transport protein sec71.
 GN SEC71 OR SEC7A OR SPAC4D7.01C OR SPAP8A3.15C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=2148401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgourou J., Peat N., Hayles J., Baker S., Baaham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Frazer A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez B.,
 RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Holzer H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revelante J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomycetes pombe.";
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: May play a role in protein transport (By similarity).
 CC -1- SIMILARITY: Contains 1 SEC7 domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
 CC or send an email to license@ebi.ac.uk).
 CC EMBL: AL117210; CAB55182.1; -
 DR EMBL: 298602; CAB11286.1; -
 DR PIR: T39252; T39252.
 DR HSSP: Q99418; 1PBV.
 DR GeneDB: SPombe; SPAC4D7.01c; -

DR InterPro: IPR000904; Sec7.
 DR Pfam: PF01369; Sec7; 1.
 DR SMART: SM00222; Sec7; 1.
 DR PROSITE: PS0190; Sec7; 1.
 KW Transport; Protein transport.
 FT DOMAIN 692 880
 SQ SEQUENCE 1811 AA; 206930 MW; 14AC990BB746804D CRC64;
 Query Match 1.6%; Score 122; DB 1; Length 1811;
 Best Local Similarity 17.4%; Pred. No. 15; Indels 594; Gaps 79;
 Matches 284; Conservative 234; Mismatches 523;
 QY 19 GTPQHTSPKSTL-EGQLMPELTKN-SPPPCNPGLVNMPSKSVHSLKPSDINKVAAT 76
 DB 325 GTPDRANSPIPOLSENKLTLSFEHRKSPDQVREAPLEBDSLEQQLRDAFL-LIRAL 383
 QY 77 GNLEIPDPGTDLEKQDTERPQVCMGV-----MTVLSDI-IRYFSP-SVMP-- 124
 DB 384 CKLSIKNIPIYEHEDYDKSQSMRSKLSLHLYHILRTYNNILSDINVKIRSPSTPTPLI 443
 QY 125 -----VC-HTKRVIPIH-----DGAEDLWIOAQELVNMKENVQLDF----- 160
 DB 444 DAVKOYICLALAKNVSHVLPVEISCEIFWILSLKNPFSELEVFTEIFPILMKR 503
 QY 161 -----QFDWTLINVFPSNASQ-----CYLCPSAQNGLAAGVDELMGVLDVLAQEV 207
 DB 504 TSSNQOKYVLNIFHRMCEPQTLIELYNYDCISGNTENIYERAILVTLISRIASOSTSD 563
 QY 208 PRAFVNLVD-----LSEVAEVSROYHGTWLSPAEP-----CNCS 242
 DB 564 PPSFVFRDQVLIDKRGFYHTLNDIPQLNASTIGSYVSHNPPYPDYQIRLKSVCILIS 623
 QY 243 EETRLAKVNMVQV 302
 DB 624 TLSS-----LFTWCNQTAPVETAKDDETESTSGEPP-----QKSKSEPPSAGINS 672
 QY 303 TLAMHLNMMPAGE---KDEPLSVKHGRPKKCPQSEFVLFYSYNSVYTLRQKPODK 359
 DB 673 TSMNVL-----SSGQALATDPSQFENKHKKQQLQEIQKFN-----KP---- 714
 QY 360 LEVREGAIECPDKPDSPTVPTSVHRLKPADINIVGALDGLSITAGNGAGSTPGNVLDVIT 419
 DB 715 -----KEGKILLASHFLASTKPTDIKAF-----LISTGLDKAVVG 751
 QY 420 QYRGSLMSVSGENIGTVT-----LANIRENPSLKGSVGTGKETSNAFL 468
 DB 752 EYLG-----EGNDENIATMHSFVDHMSFNDIPRVNMLRSP-----LQKR----- 791
 QY 469 NQAVAGRAEDLPVQARRVLDLMDKDTIRHPQEDWKIITLFIGNDLCOFCDNLVHYSPQ 528
 DB 792 -----LPGEKQIDRFM-----LKFARK 809
 QY 529 NPTDNIGKALDILHAEVPPAFVNLVTVLEIVNLRELYOKKVCPRMILRSLCPVLAKEP 588
 DB 810 YIDDNLG-----VFNQADPAILAVSIIML-----NTDLSHPQVKNRMTCQDPIKKN 856
 QY 589 ---DNSTELAT--LIEPNKKPQEKTHQLESRGYRTREDFTVAVQPFENVDMPKTSSEGL 643
 DB 857 RGVDDGANSIDSEFLVEVEEIOK--NEIVLKEOOPFTSNFP-----EIPGTS--- 901
 QY 644 PINSFPAP-----DCPHFSKSHSRAASALNMMLLEPVQK----- 679
 DB 902 --NSLSPANTISALATVGRDLQREAVYMASNMANKTEALFYDLIREQREKLSGNDIY 959
 QY 680 -TTRHKFENKINITCP--NOVQPLRTYKNSMOGHTWLPQRDRA----- 721
 DB 960 YIARR-FEH-----VCPMEFAVMPIILAAFESELDQ-----LSSDPLILQSLDGRILANN 1008
 QY 722 -----PSALHPTSVHARLPADIQVVAALGDSLTAGNGIG 755
 DB 1009 VIFPFSMDLPRAVAFQUTLKFTHLNNSTELKVTNNHALKTL-LETISLHAGDKLR-----D 1062
 QY 756 SKPDDLPTVTQYRGSLYSAGSD-GSLENVITTLPIILIEFNNLNGVAVGTODANDTNAF 814

```

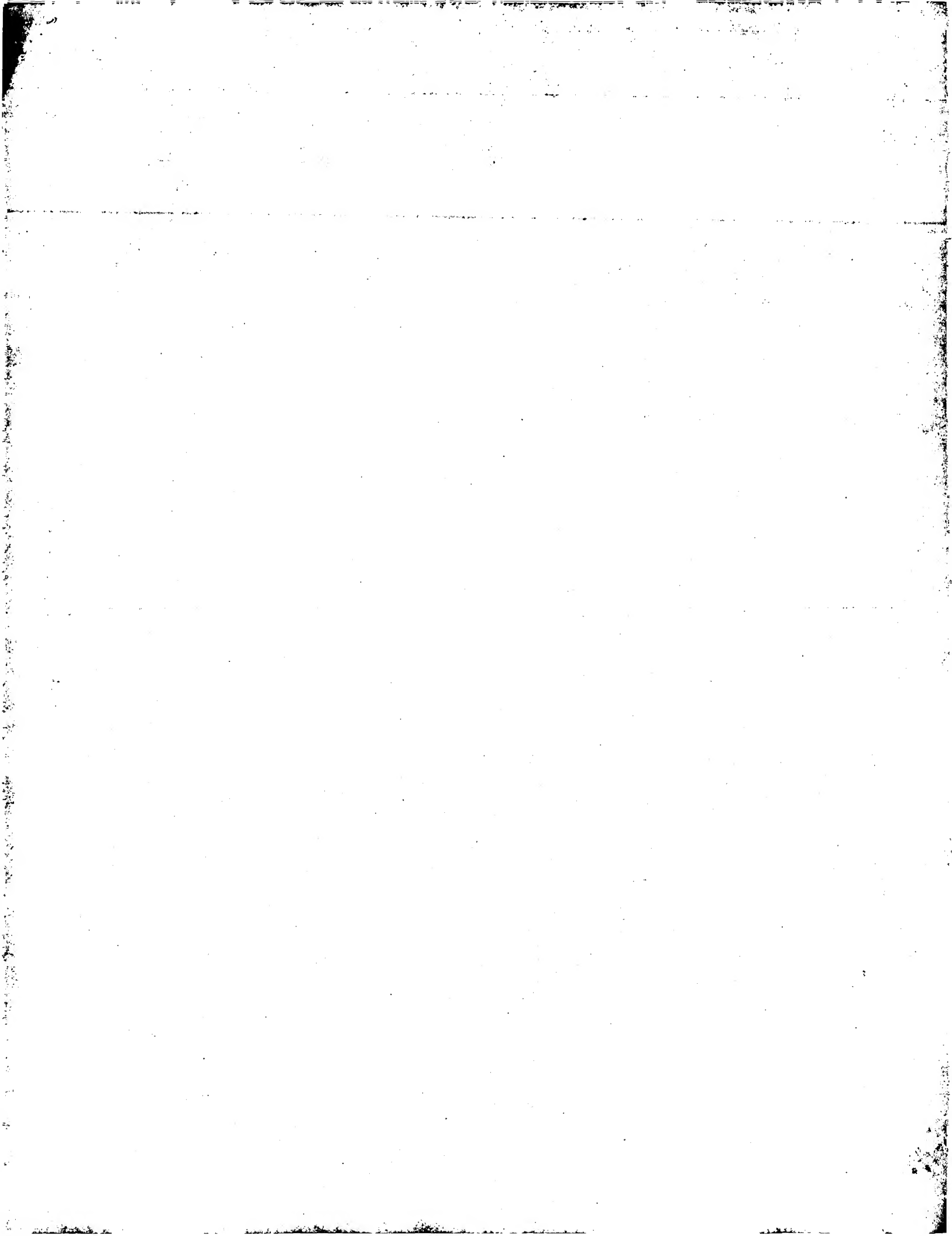
Db 1063 SMKDVLLICISQLERVQLISAGVDINSIPDVSTTKPLKSLDKNIRGRSGSISLKHKSXF 1122
Qy 815 LNVQAVPAKADLMDSQVOTLMQKKDKDRNRNFHEDMAYIVLVIGSLDCLCYCDSDNLYXA 874
Db 1123 QASVSTHSKTS-----SSVEIVRE-----YSSRRVNAV-----DML-FSNTRNLGSE 1163
Qy 875 ANFDHLRNLADVLHREV-----PRV-LVVLVD-FLNPTIMROV-----FL 913
Db 1164 GIY-DFPKALIEVSWMEIEGSLSNRRLPSLOLVLEISTYN--MRIMEMSSISL 1219
Qy 914 GNP-DKCPVOQASVLCNVL-TIRENSQELARLEAFSRAVRSNRELVSGRYDTQEDFS 971
Db 1220 GTFTVQVSCHENSITIASFALDSLRFQSMQFLIEELS-----HFKFQDF- 1264
Qy 972 VLVLPF---FONQLPVLAGLPTSPFAPDCIHPNCKHSQLARALMTNMLEL--GSK 1026
Db 1265 --LQPFHAMENSODLKIKD-----LVLRCDQMKARYONIRSGMRTI FHLIAYASK 1315
Qy 1027 TETLDLAEMLPTCTPTQNEPFLRTPRNSNTYPIKPAIENWGSDFLCTEWKASNSVPTSV 1086
Db 1316 IENL-----LVLOCA----- 1325
Qy 1087 HQLRPADIKYVAALGDSLTTAV---GARPNNSDLPFSWGLSWSIGDGNLEHTHTLPN 1143
Db 1326 -----ISVSSIGHHEHISCVLTOGAYIDLISCI-TKPAKL-----NGNOKFCISCDV 1371
Qy 1144 ILKKEN-----PYLGFSTSTWEGTAGLVAAAGAR--- 1174
Db 1372 MLKLEHELLIKHLKMKKESVYSKKLEEWYLPPLSPFNELIC-ASDLEVRSAKLVLP 1430
Qy 1175 -----ARDMPAQAMDLYER-----MKSPIDINEKD-----WGLVTLFGVN 1211
Db 1431 DCLYRHADDDEERWETVNSKALLISTISITNSQILYAKRTRETEVWMLTMV--- 1486
Qy 1212 DLCHYCNPPRAHLATEVYQHIOALDILSELPRAFVNVVEWELASLYOGQGGCAML 1271
Db 1487 -----BALKA-FIELIKNLFERLHFLPRA-INTLE----- 1515
Qy 1272 AQNNCTCLRHSGS-----LEKQELKVNNNLHGHSFSTWYHQT---QRED 1316
Db 1516 ---KCIQENSMISKVGLSCFSQFVLKNKQKQVDM--EINISINQLQMLTPIELRD 1570
Qy 1317 FAVVVPQFFQNTLTPLNBERG--DTDLTFESEDCHFSDRGHAB-----MALALW 1363
Db 1571 PSLYPOVNSSSLDEVENSFRPHIEIRFNSQSVFKSKKHILKSIYVKTQLQMLMNCW 1630
Qy 1364 -----NNMLEPVGK 1373
Db 1631 ELFHSNMLTNIPKR 1645

```

```

RN [2]
SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10711132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gccayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Chame M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikros G.L.G.,
RA Abell J.F., Agbayani A., An H.-O., Andrews-Plambeck C., Baldwin D.,
RA Baillew R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Bocham M.R., Bouck J., Brokstein P., Broctier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobery C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclet J.M.,
RA Palazolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Rainert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirbas R., Tector A., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Waasman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [3]
FUNCTION.
RX MEDLINE=91092252; PubMed=2124970;
RA Keith F.J., Gay N.J.;
RT "The Drosophila membrane receptor Toll can function to promote
cellular adhesion."
RL EMO J. 9:4299-4306(1990).
CC -1- FUNCTION: REQUIRED FOR DORSAL-VENTRAL EMBRYONIC POLARITY. MAY
ADHESION.
CC -1- FUNCTION AS A MEMBRANE RECEPTOR. PROMOTES HETEROPHILIC CELLULAR
ADHESION.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC -1- SIMILARITY: Contains 21 leucine-rich (LRR) repeats.
CC -1- SIMILARITY: Contains 1 TIR domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL, M19969; AAA28941.1; -
CC EMBL, AE003758; AAP56624.1; -
CC PIR, A29943; A29943.
CC HSP, O60603; 1FYW.
CC FlyBase, FBgn0003717; TL.
CC GO, GO:0004888; F:transmembrane receptor activity; NAS.
CC GO, GO:0006963; P:antibacterial polypeptide induction; IMP.
CC GO, GO:0006966; P:antifungal humoral response (sensu Inverteb. .; IMP.

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2004, 14:46:04 ; Search time 44 Seconds

(without alignments)
3186.680 Million cell updates/sec

Title: US-10-054-691-2

Perfect score: 7766
Sequence: 1 MGRPGIFLELLILIGQT.....RCRRGRREDPPWRLTVAL 1458

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5987.5	77.1	1458	2 A45665	adult-specific brn
2	613.5	7.9	382	2 T20655	hypothetical prote
3	605.5	7.8	414	2 T26083	hypothetical prote
4	463.5	6.0	981	2 T16060	hypothetical prote
5	406.5	5.2	348	2 T24016	hypothetical prote
6	347	4.5	349	2 T24015	hypothetical prote
7	338	4.4	425	2 T21835	hypothetical prote
8	159.5	2.1	941	2 F97353	uncharacterized co
9	149	1.9	3655	2 T38084	TRAP-1-like protein
10	146.5	1.9	1755	2 F82618	chemotaxis-related
11	141.5	1.8	913	2 T15278	hypothetical prote
12	133	1.7	1708	2 F69189	protoporphyrin IX
13	131.5	1.7	1361	2 T30884	neural specific DN
14	130.5	1.7	891	2 E96590	hypothetical prote
15	130.5	1.7	1165	2 T15279	hypothetical prote
16	130.5	1.7	1545	2 T26589	hypothetical prote
17	129.5	1.7	1245	2 T42920	hypothetical prote
18	128	1.6	589	2 G86550	polysaccharide bio
19	127.5	1.6	1016	2 T30553	disease resistance
20	127	1.6	1280	2 T34357	hypothetical prote
21	126.5	1.6	1498	2 E86302	hypothetical prote
22	126.5	1.6	2471	2 T42977	large tegument pro
23	126	1.6	1028	2 E51173	myosin I beta - bu
24	125	1.6	1178	2 S78475	mannosylphosphoryl
25	125	1.6	1650	2 S28721	hypothetical prote
26	124	1.6	1946	2 AC2141	serine/threonine k
27	123.5	1.6	1230	2 E64664	outer membrane pro
28	123.5	1.6	1534	2 G91017	probable membrane
29	123.5	1.6	1534	2 A85862	hypothetical prote

30	123.5	1.6	2672	2 A48126	translation activa
31	122	1.6	1811	2 T39252	probable protein t
32	121.5	1.6	1097	2 A29943	Toll protein precu
33	121	1.6	4725	1 A44357	dynein heavy chain
34	120.5	1.6	836	2 D97182	extracellular neut
35	120.5	1.6	1570	2 T38792	probable protein t
36	120.5	1.6	2843	1 RBHUP	adenomatous polyo
37	120	1.5	1237	2 D71850	probable outer mem
38	120	1.5	2140	2 P95074	serine proteinase,
39	120	1.5	15281	2 S41309	cyclosporin synth
40	119.5	1.5	1163	2 E71481	probable swf/bnf h
41	119.5	1.5	1276	2 S11455	botulinum neurotox
42	119	1.5	1107	2 T25450	hypothetical prote
43	119	1.5	1116	2 T16112	hypothetical prote
44	119	1.5	2089	2 C85426	ATM-like protein l
45	119	1.5	3738	2 T05501	hypothetical prote

ALIGNMENTS

RESULT 1

A45665 adult-specific brush border esterase/phospholipase (EC 3.-.-.-) precursor - rabbit

C/Species: *Oryctolagus cuniculus* (domestic rabbit)

C/Date: 03-May-1994 #sequence_revision 03-May-1994 #ext_change 08-Oct-1999

C/Accession: A45665

R/Bol1, W.; Schmid-Chanda, T.; Semenza, G.; Mantel, N.

J. Biol. Chem. 268, 12901-12911, 1993

A/Title: Messenger RNA expressed in intestine of adult but not baby rabbits. Isolation c

A/Reference number: A45665, MID:93286138, PMID:8509424

A/Accession: A45665

A/Status: Preliminary

A/Molecule type: mRNA

A/Residues: 1-1458 <BOU>

A/Cross-references: GB:Z12841, NID:91689; PIDN:CAA78303.1; PID:91690

C/Keywords: hydrolase

Query Match	77.1%;	Score 5987.5;	DB 2;	Length 1458;
Best Local Similarity	77.6%;	Pred. No. 0;		
Matches 1130;	Conservative 127;	Mismatches 197;	Indels 3;	Gaps 3;
QY	1	MGRPGIFLELLILIGQTPOIHNSPKSTIEGOWPBTLNKSPPCPNKLGUMPSK 60		
DB	1	MLMPSPVFLGLPLGRGADQIQSSGGNTLEGOLMPBSLTFPPCPKTLAEVPSR 60		
QY	61	SVSLKPSDIKFFVAIAIGNLEIPDPDGTGLEKODWTE-RPOOVCMGVMTVLSDIIRYFSP 119		
DB	61	SVHSLRPSDIKFFVAIAIGNVETAPDSGADDLDEODGTETKRPDQACGVTVLSDIIRFSP 120		
QY	120	SVPMPCVHTGKVIIPHDGAEDLWIOAQLVNNKENTLQDFQDWKLINVFSNMQCYL 179		
DB	121	SALMPLC-DETRLVPRGCAEDLWMOQTETLVSRBNPQDFEHDKLVNVFSNTSQCFP 179		
QY	180	CPSAQNGLAAGVDLMVLDYLOQEVPRAYNVLDSEVAEVSQYHGTMLSPAPERC 239		
DB	180	CPSAQNGKVLGMDLRTLLDYLOQEVKAFVNLVLSBELAFAFRMGQALSPAPERC 239		
QY	240	NQSEETRLAKVVMQWSVGEAANSLASGRYSEOSFTVPQFPFETPSLHSEDPRLQ 299		
DB	240	RCIRETSQTLTKVLTQWSTYEAANDSLASSKVTQBSFAVVFQFFESSLSALLAEPRLQ 299		
QY	300	DSITLAMLNMMMPAGEKDEPLSVKHGRPMKCPQSQSPYLSYNSNYLTRLQRPDK 359		
DB	300	DETTLALSLNMMMPBEGKKEPFSEKERRPLRCPTQESPYLFTYNSQQLRVSPQCK 359		
QY	360	LEVRECAEIRCPDXPSDVPVPSVRLKPADINVIGALDSITLAKGAGSTGNVLDVLT 419		
DB	360	LEVRETEIRCPDXKPSDVPVPSVRLKPADIKVIGAMDSTLAKGAGSOGNTLDVLT 419		
QY	420	OVRGLSMVSGDENIGTVTLTANILREPNPSLKGSVVGKETSPPAFNQAIVAGRAD 479		

Dh	420	QYRGISMSVGGDQNI	STVTTTLANTILARENP	SLQGRSVGTGR	ETTSGAF	AFNOAVAGARADG	479									
Qy	480	LPVQARRLVDLMK	QDTRIRHFOEDMKIT	TLFGNDL	CDPCNDL	VHYSPQNTDNI	539									
Dh	480	LIPARQIRVALAMK	QDTRINFQEDMKIT	IVTFGNDL	CDPCNDP	PVRYSPQNTDNI	539									
Qy	540	ILHAEPRAFNVL	VTVEIVNLRELJOEK	YVCPMILRSLC	CPYLKCPDDN	STELATILE	599									
Dh	540	ILHAEPRAFNVL	KVLEISIKLRELJO	ETQTKVSCPMILR	SLCPCYLKCPDDN	STELASILE	599									
Qy	600	FNKKFOEKTHOL	IESGRYDREDFT	VVVOFPFENV	DMKTSBGL	PDNSFFAPDC	HFSSK	659								
Dh	600	TIKEYOERTQO	LIDSGRDTRED	DFEVLVQPF	EKKVMKPTOD	GLPDNSFFA	PDCHFSSK	659								
Qy	660	SHSRAASALMM	NMLEPVQOKT	TRHKFEKKNIT	CGNOVQF	LRTKYMSOGH	TMLPCRD	719								
Dh	660	AHAHAASALMM	NMLEPVQOKT	THNDEBSA	VNITCGNOV	PEPLSTKYNSVQ	GGTMLPCRD	719								
Qy	720	RAPALHPTS	VHALRPADIQV	VAALGDSL	LTAGNGIS	KPRDLPVTTQY	RGSYAGDG	779								
Dh	720	RSPBASPT	SHALRPADIQV	VAALGDSL	LTAGIGIS	KPRDLSOD	TTQYRGISYSSG	779								
Qy	780	SLEAVTTLP	ILREBPNML	TGYAVGTG	DANDTNAF	LNOAVGAKA	EDIMSOYTLMQ	839								
Dh	780	SLEAVTTLP	ILREBPNML	TGYAVGTG	DANDTNAF	LNOAVGAKA	EDIMSOYTLMQ	839								
Qy	840	DDHRVNP	FEDMKVITV	LIGASDL	CDYCPDS	NLYSAANVDH	RLNMLDYAREV	PVNL	899							
Dh	840	DDHRVNP	FEDMKVITV	IGASDL	CDYCTDS	NLYSAANVDH	RLNMLDYAREV	PVNL	899							
Qy	900	VDLNP	PTIMRQV	FLGNPD	KPCVQOASV	LNCVLT	IRENSOELIARE	SRAVYRSSMBELV	959							
Dh	900	VDLNP	PTIMRQV	FLGNPD	KPCVQOASV	LNCVLT	IRENSOELIARE	SRAVYRSSMBELV	959							
Qy	960	GSGGYDPO	EDSSVLOPF	FNQNL	OLPYLAD	GLPDTSF	APDCLHPQKH	SOLARLYMNM	1019							
Dh	960	ESGRDTR	EDSSVLOPF	FHSIQ	LPVLQD	GRDITFS	FAPDCVHPQKH	SOLSRALYMN	1019							
Qy	1020	LEPIGSK	ETEDDLRA	EMPIIT	CTPQNE	PLRTPR	NSNTYYP	PIKPALEMMG	SDFLCTEMKRS	1079						
Dh	1020	LEPIGSK	ETEDDLRA	EMPIIT	CTPQNE	PLRTPR	NSNTYYP	PIKPALEMMG	SDFLCTEMKRS	1079						
Qy	1080	NSVPTSV	HOLRPADIK	VVAALG	DSITTA	VGARPNNS	SDLPTS	WRGLSMSIG	DGNLETHT	1139						
Dh	1080	RGVNP	SVHLEOP	CGDIKV	VAALG	DSITTA	LGMGARPNNS	SDPPEMF	WRGLSMSIG	DGNLETHT	1139					
Qy	1140	TLPNIL	IKKENY	YLIGFS	STWEGTA	GLNVAE	GAARBM	PAQAMDL	VBERMKS	PINILEK	1199					
Dh	1140	TLPNIL	IKKENY	YLIGFS	STWEGTA	GLNVAE	GAARBM	PAQAMDL	VBERMKS	PINILEK	1199					
Qy	1200	DMKLV	TLF	IGVNDL	CHYCE	NPEAH	LATEVYOH	IQOALD	ILSEEL	PRAEVNVE	WEMLASL	1259				
Dh	1200	DMKLV	TLF	IGVNDL	CHYCE	NPEAH	LATEVYOH	IQOALD	ILSEEL	PRAEVNVE	WEMLASL	1259				
Qy	1260	YQGGG	GKKA-M	LAAONN	CTCLRH	SGSSLE	KQELK	VYMN	LGHG	ISSBSY	WQYTORE	FEPA	1318			
Dh	1260	HODQGG	RCATILAA	QSHCT	CKEYSSQ	SVEMOEL	KKVNN	LNLOS	GLSR	LSYS	HQYORE	FEPA	1319			
Qy	1319	VVVO	PFCNTL	TLPLNER	GDJTL	TFPSE	DC	CPHPS	DGHA	EMALM	ANNML	LEPV	GRKTSNN	1378		
Dh	1320	VVVO	PFCNTL	TLPLNER	GDJTL	TFPSE	DC	CPHPS	DGHA	EMALM	ANNML	LEPV	GRKTSNN	1379		
Qy	1379	FTHS	RAKLK	CPSP	PS	PYL	YTLRNS	HLBP	DOAE	AE	EVLYMA	PVAAGV	GLVVG	1438		
Dh	1380	FTY	SR	TKL	K	CPSP	PS	PYL	YTLRNS	HLBP	DOAE	AE	EVLYMA	PVAAGV	GLVVG	1439
Qy	1439	RCR	GR	GR	ED	PP	MS	LT	1455							
Dh	1440	RG	MC	R	P	E	D	P	P	L	S	T	1456			

hypotheicalprotein F09C8.1 - Caenorhabditis elegans
 C|Species: Caenorhabditis elegans
 C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
 C|Accession: T20655
 R|McMurray, A.
 submitted to the EMBL Data Library, November 1995
 A|Reference number: Z19306
 A|Accession: T20655
 A|Status: preliminary; translated from GB/EMBL/DBJ
 A|Molecule type: DNA
 A|Residues: 1-382 <Wt>
 A|Cross-references: EMBL:Z68132; PIDN:CAA92221.1; GSPDB:GN00028; CESP:F09C8.1
 A|Experimental source: clone F09C8
 C|Genetics:
 A|Gene: CESP.F09C8.1
 A|Map position: X
 A|Intons: 7/1, 71/3, 109/1, 164/3, 192/3, 245/2, 263/3, 303/3, 360/3
 C|Superfamily: Caenorhabditis elegans hypotheical protein R07B7.8

[illegible]

RESULT 3
 T26083
 hypothetical protein W02B12.1 - *Caenorhabditis elegans*
 C1Species: *Caenorhabditis elegans*
 C1Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Mar-2002
 C1Accession: T26083
 R1Swingburne, J.; Ainscough, R.
 submitted to the EMBL Data Library, October 1995
 A1Reference number: Z20149
 A1Accession: T26083
 A1Status: preliminary; translated from GB/EMBL/DBD
 A1Molecule type: DNA
 A1Residues: 1-414 <WII>
 A1Cross-references: EMBL:Z66521; PIDN:CAA91393.1; GSPDB:GN00020; CESP:W02B12.1
 A1Experimental source: clone W02B12
 C1GeneticB:
 A1Gene: CESP:W02B12.1
 A1Map position: 2
 A1Introns: 5/1; 6/3; 101/1; 156/3; 212/2; 273/2; 297/3; 332/3; 368/3
 C1Superfamily: *Caenorhabditis elegans* hypothetical protein R07B7.8
 Query Match 7.8%; Score 605.5; DB 2; Length 414;
 Best Local Similarity 36.8%; Pred. NO. 1.9e-33;
 Matches 141; Conservative 57; Mismatches 118; Indels 67; Gaps 10

```

QY      376 SDVPSPSARLKPADINVIAGLSDSLITLNGAGSTPGVAVLVLTQYRGSLMSVGSDENIG 435
Db      35 SKVPPSSASASVRADIKVIGLGDSTITANGAGAPGDPPLVITIQYRGSLAQIQGGSDSL 94
QY      436 TVTTLANILIREFNPSLIGFVSVGKETS- PNAFLNQAVAGRAEDLPVQARLVTLMDKND 494
Db      95 EHITVANLRKFNRPVLVGASKGIGSENVWEVSHLMVGVRGASEKDIIQARALVTMHAI 154
QY      495 TRHHPEDMKITLTFPGGNDLDFC-----NDLVHYSQNTDNIIGKALDILAEV-- 546
Db      155 SEINVKEDKLVNIFIGANDICVYCEDPYFNSTALHGNATEEKIIIAVOQLDQNDPRRT 214
QY      547 -----RAFVN-----LVTLVLEVNREL--VQEKVCVCPRIILR 578
Db      215 NPKFKLPSRLSVCKTFSSWREKNSNIYSRTVSLTGFPNRMRLRKDIKKKYFEGGLHTF 274
QY      579 SLCPVLKEDDNSTELATLIEFNKKFQEKTHQ-----LIESGRYDREDFTV 626
Db      275 E-CDC-----ESNQKQPTDDDIQGVCGYMAEKDIQNTGLGLDNMDPTFV 318
QY      627 VQPFEEV-DMPKTSGELPDSFPAADCFHSSKSHSPASALNNMLERPVGQKTTIRHKF 685
Db      319 VQPFEEGLIDPYPASGVVDMTFPAADCFHSGAIGHGNIQMLNMTIVQPVGKQISVNL 378
QY      686 EN-KINITCPNQVQPLRTYKNS 707
Db      379 SDPSVGLHCPSINCPFPPTKNS 401

```

```

RESULT 4
T16060
hypothetical protein F13H8.5 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C/Accession: T16060
R/Ding, H.
submitted to the EMBL Data Library, July 1995
A/Description: The sequence of C. elegans cosmid F13H8.
A/Reference number: Z18455
A/Accession: T16060
A/Status: preliminary; translated from GB/EMBL/DBD
A/Molecule type: DNA
A/Residues: 1-981 <DN>
A/Cross-references: EMBL:U02119; NID:G722370; PID:G722375; PIDN:NA046681.1; CESP:F13H8.5
A/Experimental source: strain Bristol N2
C/Genetics:
A/Gene: CESP.F13H8.5
A/Intons: 46/1, 124/1, 201/3, 255/2, 419/1, 452/1, 645/3, 666/3, 692/3, 719/3, 794/3, 8

```

[illegible]

QY 653 CFFHSSKSHSAAALNNMMLPEVQOKTRAKREKINI-----TCPNQVCEPLRYKKS 707
 Db 326 CFHSLSPAHDLAAQIKMGKLEPEIDQKI-----ITNQSLVGDFRVCPCPEECYLYTIQNS 381
 QY 708 MCGHGHTWLPCHD---RAPSAHLRTSVAAALRPADIQY 740
 Db 382 ENCE-----PSREMRFLRPVSPASSPGSLSPMYKILPV 414

RESULT 5
T24016
hypothetical protein R07B7.8 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
C/Accession: T24016
R/Harris, B.
submitted to the EMBL Data Library, July 1996
A/Reference number: Z19830
A/Accession: T24016
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-348 <WIL>
A/Cross-references: EMBL:Z75955; PIDD: CAB00118.1; GSPDB: GN00023; CESP: R07B7.8
A/Experimental source: Clone R07B7
C/Genetics:
A/Gene: CESP:R07B7.8
A/Map position: 5
A/Intons: 4/1; 55/3; 91/1; 147/3; 223/2; 277/3; 335/3
C/Superfamily: *Caenorhabditis elegans* hypothetical protein R07B7.8

Query Match	5.2%;	Score	406.5;	DB	2;	Length	348;
Best Local Similarity	32.2%;	Pred. No.	5.9e-20;				
Matches	117;	Conservative	66;	Mismatches	133;	Indels	47;
						Gaps	11.

QY IKALEIEMWGSDFLCTEWKASNSV-----PTSVHQLRPADIKVVAALGDSLTVAVGARN 1113

Db 15 VKMALE-----KYNISDPHEIYDDDVNMKPHHIRIVAGMGDSL--IGSRAE 61

QY 1114 NSSDLPTSWGSLWSIGGGDNLETHTTLPNLK--KPNPYLIGFSTSTWEG--TAGLVN 1168

Db 62 NIVGQRQRYPGNAFFPTGMDFEVDRLHLLTVVNIIFRIAEKTGNKLGSGSTGIDYGENTGLNV 121

QY 1169 AAEARARMDPAQAWDVERMKSSPDNLNEKDWLTLTLFTGVND---LCHCENPEAHLA 1225S

Db 122 AIGGMSKDDILIRAKELVSRISKANKEININEMDWLSLWIGTDVGTGLGYRLDPPIR--V 179

QY 1226 TEVYOHFOQALDLISELPFAFNVVVEWELASLYOOGGKGCMALAAQNNCTCRHSOSS 1285S

Db 180 DEYKSHIEKGLYKENLPRTIVSIYGMFPAQLIQEAQ-----SILKNGKRA 226

QY 1286 LEKOELKKNWNLQHGSISSFSYWHQYQR--EDFAVVVQPF-FQNTLPLYNERGDTLT 1341

Db 227 RLYENQKKLD-DLSDGYRNVSYDQNNHEHNSNDFTLVVGPRATEYIDSYRDEHGKKNPT 285

QY 1342 FFSFEDCHPSDRGHAEMALALMNNMLPEPVGRKTTSNPFHSRAKLKCPSPESPVLVTLRN 1401

Db 286 FYASDLFHLISKFGHAVALAKHYWNLPEPVGEKTRADLGCTKPKXIYELNEXNCLIKTVGN 345

QY 1402 SRL 1404

Db 346 SKM 348

RESULT 6

T24015

hypotheical protein R07B7.9 - *Caenorhabditis elegans*

CiSpecies: *Caenorhabditis elegans*

CiDate: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000

CiAccession: T24015

RiHarris, B.

submitted to the EMBL Data Library, July 1996

A:Reference number: Z19830

A:Accession: T24015

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1349 <Mtu>
A:Cross-references: EMBL:Z75955; PIDN:CAB00117.1; GSEPB:GN00023; CESP:R07B7.9
A:Experimental source: clone R07B7
C:Genetic8:
A:Gene: CESP:R07B7.9
A:Map position: 5
A:Introns: 4/1: 55/3; 92/1; 148/3; 224/2; 278/3; 336/3
C:Superfamily: Caenorhabditis elegans hypothetical protein R07B7.8

Query Match	4.5%	Score 347	DB 2	Length 349
Best Local Similarity	31.9%	Pred. No.	7e-16	
Matches 106	Conservative 52	Mismatches 124	Indels 50	Gaps 11

```

OY 374 PPSDVPVPSVYRHLKXADINVLGALDSSLLTAGGA-----GSPNGVLDVLTQYRGIS 425
Db 27 DPRETYPPDVNSVKKPHHLRVIGAMSDSLTGYCASHFTEIRLNGPNPN-----S 75
OY 426 MSVGDEINIGVTTLANILR-----EENSLKFSVGTGKETSPPAFLNOAVGGAEDLP 481
Db 76 FFTGDEIEITDGLSLSTYINFRVIAETGNKLFGSGSGVG--YANNTGVLNAVCGKMSDDL 133
OY 482 VQARRLYDLMKNDPTFIHFQEDWKIITLTFISGNDLDCFNDLVHASP-----QNFPTDNGK 536
Db 134 RQAKDLVSRIRANKENINEXKOMKVLWSIMGTNDV---GNLVFSGENPIPVKEYAFIEE 189
OY 537 ALDILIAEVPRAVNLVTLEIVNLRELYOEKKVCYCPMILRSLCPVLKFDNDSTELAT 596
Db 190 GLTYLKKKLLPRTVSIIGMFPOLLQENAV-----YLRTGNRPGRP-----ESRKLDE 238
OY 597 LIENFKKQOEKTHOLIESGRVYTRDFVYVQPF--FENVDMKXTEBGLPDSFPAPDCGH 655
Db 239 LCD---SYRNASYELQNEGKFPDRE--FVYVQPFTEYTDAPRNEFGYNSALYAVDFH 294
OY 656 FSSKSHSPASALMNNMLPEVGOKTRHAFEN 687
Db 295 LKRLGOALVAKHYWONLFEVGEKTKYRDELGN 326

```

RESULT 7
T21835
hypothetical protein F36A2.9 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21835
R:Lennard, N.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19476
A:Accession: T21835
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-425 <EWL>
A:Cross-references: EMBL:Z81077; PIDB:CAB03072.1; GSPDB:GN00019; CESP:F36A2.9
A:Experimental source: clone F36A2
C:Genetics:
A:Gene: CESP:F36A2.9
A:Map position: 1
A:introns: 90/2, 125/2, 196/1, 267/3, 316/2, 354/3

[illegible]

```

Db      243 RLKLSLSDYNNEMITITIGTEELSCRCBPSI-----DNIRKALIEHQIETPKA- 294
QY      550 VNLVTVEIYNLRELQOEKRYVCPMILRLSLCPVYLKPEDNSTELATLIEFNKKFOE-- 606
        |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      295 --LVVLLGPVAVSSFHEQKS-----NLKAKACAC-----SRDTE-GFMVDVRKMSKVMR 342
QY      607 KTHOLIESGRYDTRDFVTVVQPPFENVDMPK-TSEGLPDNSFFAPDCPFHSSKSHRAA 665
        :|||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      343 DVQKFEVNG-----VTARPTGMISYPMVTTTSRYPSQLFTRDKFLNRRGHNVAT 393
QY      666 SALNNMML 673
        |||||:::|
Db      394 KMLNNRLI 401

```

RESULT 8
 PF7353
 Uncharacterized conserved membrane protein, YVEB_B. subtilis homolog CAC3696 [imported]
 C:Species: Clostridium acetobutylicum
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C:Accession: PF7353
 R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: PF7353
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-941 <KUR>
 A:Cross-references: GB:AE001437; PIDN:AAK81617.1; PID:g15026801; GSPDB:GN00168
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC3696

Query Match	2.1%	Score 159.5	DB 2	Length 941
Best Local Similarity	18.1%	Pred. No. 0.027		
Matches 178	Conservative 143	Mismatches 343	Indels 321	Gaps 45

```

0Y      412  GNVLVDLTVQRYGSLMSVCGD--ENIGATYTTLANILREFPNPSLAKGSVGKSTSPAPFLN  469
      30  GVILFVLVCAVAGLITFLNRNHHKQVQVOTTKPIALVN-----EDQPTFNG  77
0Y      470  QAVAGRAEDLPVQARLVDLMKDTRTIFQEDWKIITLFIGNDLCEPCNDLWHYSPON  529
      78  TNYNFGT-----FVNLVSDNKK--NMQVLSRSVANKKAYADSDVAVIYLPQN  124
0Y      530  FTDNIGKRLDILHAEPFVNLVTVLIVNRELJOEKYVCPRMILSLCPCVLKPD  589
      125  FTHNI-----LTLQALNPQKELDYKVLASKSOLN-----  154
0Y      590  NSTELATLIEFNKKFOEKTHOLIESGRYDTRDFVWVVOPEFPE-----NYDMKTSGLBD  645
      155  -----NKLLQDKIVNVL-----YDFN--TSIVGMYYASVAGVANVAQTUNWG--  193
0Y      646  NSFAPDCFHFSKSHSRALAMNMNLEPVGQKTKTRHKFEKINITGCNOVQPFRLTYK  705
      194  -----NVVNSQGTLLSSLDNNVGPF--QTTNBSYSSVSIA-----  228
0Y      706  NSMOG-HGTWLPCHDRAPSALHPTSVHARPADIQVVALGDSLTPAGNGISGKPDLDPV  764
      229  NGLASENSDWIQEONSPTNS-----VVNLNBSSSSPNG-----  263
0Y      765  TTQYRGLSYSAGGDSLENTTLPILIEFN--RULTYAVGTGANDTNALNQAVPBA  822
      264  -----LPLNTYFPTQKIT-----DINLTNA--NOGITNQ  292
0Y      823  KAED-----LMSOVOTLQMKMKMDHVFNFHEDKVITVLIGSDLCYCTDSNLSYSA  875
      293  ANSKSKYYDDQYTAAYKAVQAVQOQFHTD-----ASGNETGYVALTKQIOISSY  341
0Y      876  NFV-----DHLRNALDVLHREVPRVLVNLVDFLNFPTMRQVFLGNBPDKCPVOQASVYLCNC  930

```

Db 342 NTLISGVHDNIGTQIOISLTGK--CTTLNMLEKMDWLRFEYGDISPTPDK----- 388
Qy 931 VLLRENSOELARLEAFSRAYSRSMELVSGRYDQEDFSVVLQPFQNIOLPVLADGL 990
Db 389 -----TDFISLETADNA--RSAMAMMIATS--FGKDNISGTAAPPKTINNLNLSNVV 438
Qy 991 PD-----TSFPAPDCIHPNCKFSQALALMTMLLEPGSKTEITLDLRAEMPITCPTONE 1045
Db 439 GDDVDTLATLVKNGSILTPQOKDYEAQLGVLNKYATAFGIKITADPSA-----APRYVN 493
Qy 1046 PLKRPNSNYTPPIKPAIENWGSDFLCTEW-----KASNSVPTSVQRLPADIKVVAALG 1101
Db 494 -----SNEFTKQITTYVKK-----ATEYTLKLQASNGLSNS-----GISIVG 531
Qy 1102 DSIITTVAGARPNSSDLPFTRSGLSIGDGNLET--HTLPLNLIKFNRYLLGFSTISW 1160
Db 532 KPTGTALVNPDSVTLDNRTTSTIDS--KGNVTTAPPTISP-----MTYVYTYKVLDG 582
Qy 1161 EGTAGLNVAA--EGARARD-----MPAQ-----WDLVERMKSPDINLEKDM 1201
Db 583 QAASGVTVASVNDGSGGTSSVENFGIIPANALSEVAGSGQFGYISTILN-----NIDTAS 638
Qy 1202 KLVTLFIGNVNDLCHYCENPEAHATEYVQHIQQA-----LDI-----LSEELPRAFY 1248
Db 639 SLITFLYGARG--ATYSIDIKSCTKVSDFQASQOSIFKMGVNDVQTQMSRLSDDDVQKFM 697
Qy 1249 NVVEWELASLYOGGCKMMLAQNCTCLRHSSQSLERKELKKNWNLOHGISSESTY 1308
Db 698 NY-----GKANIKNVITDITLTLLNSINIASIOKEKA-----TLSDNLI,PS--NYF 737
Qy 1309 HOYTQREDPAVVVQPFQONTLPLN 1333
Db 738 SOVSSD-----LQAWYNQMTDLTN 756

RESULT 9
TRAP-like protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T38084
R:Genes: S.; Odeili, C.; Church, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z21768
A:Accession: T38084
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-3655 <GEN>
A:Experimental source: strain 972h-; cosmid c1f5
A:Genetic: SPDB:SPAC1F5.11c
A:Gene: SPDB:SPAC1F5.11c
A:Map position: 1
A:Introns: 22/1

Query Match 1.9%; Score 149; DB 2; Length 3655;
Best Local Similarity 20.0%; Pred. No. 1.5;
Matches 167; Conservative 112; Mismatches 307; Indels 248; Gaps 40;
Qy 672 MLEPVQOKTR--HKFKNKINITCPNOVQPLRTYKNSM-----OGHGWTMLPC 717
Db 367 LARPLAFSTLADLLHVRDELNET---QIKSIMYSTNMHDLTLISGLQTMGARLLIM 423
Qy 718 RDR---APSAHPTSVHARPADIOVVALGSLTAGNGISKRPDDLPTVTOYRGISYS 774
Db 424 VDRMISLPSI-----PDALIFLLSIFDSFV-----NKFSELNLSLDGPFKKXYE 467
Qy 775 AGDGSLENTVTPILNLEFNRNLGVAVTGANDTNALNQAQVCAKAKEDLMSOVQTL 834
Db 468 -----EELKETKSPTRSPRDLSSSTSVN---GSFL-----FKNLMLFGIRAL 508
Qy 835 MQMKODHRVNFEDWKVITVLIGGSDLCDYCTDSNLYSAANFVHLRNLAVLHREVR 894

Db 509 MYGLR-----TCKSRCEIGCEQFSGFLTNIXKFEAVTF-----QKLFEEVGK 551
Qy 895 VLVNVLDFLNPITMRQVFLNPKCPVQASVLCNCLTREMSELARLEAFSRAIRS 954
Db 552 GF---SYRFP---EGVYLETFFCC--EESLDRPALISTLPRKDEDCLEAVATIF-- 599
Qy 955 MRELVSGRVDTQEDFSVVLQPFQNIOLPVLADGLPD--TSFPAPDCIHPNQ---KPH 1008
Db 600 -----IHLESIFLKVETN--LPTFPDQKUNTLTHIHPFLSNESTSKFL 646
Qy 1009 SOLARALMTMLLEPGSKTEITLDLRAEMPJ-----TCPONEPFLR----- 1049
Db 647 NILRFL--LSRIELGSS-----DIRGSLVLRFLRSLFVTVSMFATENEVLRYVEEI 701
Qy 1050 -----TPRNSNTYPIKPAIENWGSDFLCTEWKASNSVPTSVQRLPADIKVVAAL 1100
Db 702 VKCMKLA PNASLNYLLRALFRGIGGGRFESLVK---EVPPLHALLEAFNSLL-- 755
Qy 1101 GDSLTTAVGARPNNSDLPFS-----WRGLSMSIGDGNLET----- 1137
Db 756 -----ISARTPEKDLTELCLTTPVRLSLPLPYMSYLMRLVMSLSKSOELVSGCLR 808
Qy 1138 --HTLPLNLIKKE-----NPYLIGFSTSTW-----EGTAGLVNAAEGARADM 1178
Db 809 TPELCNDLTPDLPIAPYIEDLNNALMSHQPLPYVYNHSHTLAKIKGLGGRKRL 868
Qy 1179 PAQAMLVVERMKSPDINLEKMKVTLTFLGVA--DLCHCENPEAHATEYVQHIQQA 1236
Db 869 ---LDRVOSLKNKSPBN--NDFTLLSLKGVQPOLH-----YQYVDEAV 910
Qy 1237 DISELPRAFYVVE-----VMEASLYOGGCKMMLAQNCTCLRHSSQSLERKEL 1291
Db 911 NLLSS--PSSDLEVKQAFYVNCISKLTYKSDATNSLSTRCTADKSKSPDRRP 968
Qy 1292 KKNWNLOHGISSESTYHOYTQREDPAVVQ-----FPQNTLPLNERGDTLTFPSE- 1345
Db 969 YSVIPSRMGRSSFT---QISDSDDEITILASATYGLFAFTYDELRE-----EAYFLEK 1021
Qy 1346 -----DCHPFD---RGHAEMALAMNNMLEPVGRKTTNNFTSRKALKC 1388
Db 1022 LAVNVIVHIDIFYAFVDIQGNHSGFTTNLQKEVI-----ISPHYAHCISEVVC 1069

RESULT 10
F82618
Chemotaxis-related protein kinase Xf1952 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F82618
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; NCBI:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: F82618
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1755 <Sim>
A:Experimental source: strain 9a5c
A:References: GB:AE004014; GB:AE003849; NID:G9107044; PIDN:AA84754.1; GSPDB:GN001.
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A.
Birtone, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carrer, D.M.; Carreir, H.
as-Neto, E.; Docena, C.; El-Dorty, H.; Facincant, A.P.; Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, M.L.; Kemper, B.L.; Kitajima, J.P.; Krieger, J.E.; Kurame, E.E.; Laigret
J.D.; Junqueira, M.L.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Henck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, R.; Rosa, A.J. de M.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Teshato, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XRI952

Query Match 1.9%; Score 146.5; DB 2; Length 1755;
 Best Local Similarity 18.7%; Pred. No. 0.63;
 Matches 245; Conservative 142; Mismatches 457; Indels 465; Gaps 55;

```

Qy 193 VDELMGLVDLQGVPAFNVLDLSEV---AEVSRQYHGTW-LSPAPCCNCEETRL 248
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 57 LDBALG---QAROEIELEVENPVDTSRLISCADYLHQGTLELLEYPAPMAAEEMERL 113
Qy 249 AKVVMQMSYQEAAMNSLLASRSYSQESFTVQ-----PEFYETTPSLHSDEPLQSTT 303
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 114 AKVMQ-----VDGVNREBAQVLMRGRTLLPYLELQNGHDIP-----I 155
Qy 304 LAWHLNRMMEPAGEKDEPLSVKHGRPMKCPQESPYLFSYRNSNYLTRLOKPODKLEVR 363
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 156 VLLPLNEIRIARQGE-----SRLENVLFAPDP-----EVR 186
Qy 364 EGAEIRCPDQPS-----DTVPSTVH---RLKPA-----DINVI 394
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 187 NATEVELDHARSSLYTGCNRQLDTPAASAVYBELRVKEALDLYLRIGDVTDLRTQITEL 246
Qy 395 GAGDGLT-AGNG-AGSTPGNVLDVLTQYRGLSMVSGDENI--GVTTLTANILREPNPS 450
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 247 SSVAADTLAMMGLKAGSL-----VLOQRDALRIYSDEKADBALLIETAGALLHYDAS 300
Qy 451 LKGSVGTGKETSPNAFLNQAVAGRAEDLPVQARLVLDLMDKDTRIHFOEDWKITLFI 510
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 LDMQVANLGAEDADVVLANA-----ERRRLDVLAEHALNFTAAECPIAFI 349
Qy 511 GGNLDLCFCNDLVYSPQNTDNIQKALDILHAEPFAPNVLTVLEIVNLRLEYQEKV 570
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 350 ESN-----WQARLLKVPGLSEVGALSTILESOAASYLEGV-----RRYI 391
Qy 571 YCPMILRSICPCVLKFPDNDSTELATLIEFNKKFOEAKTHQIIESGRYDTRDFVVOVF 630
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 392 EC-ELIKRVRPARQDLTADMASLEFYEAIRE-----RRGRREDILIEIRNS 441
Qy 631 FENVDD--MPKTSGLPDNSFPAPDCFHFSKSHSRAASALMNMMLPVQOKTTRHKEEN 687
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 442 LEALSYMPIR-AEGRDVSQSTQSTDLTFSAFD-----BALPELVVEDGVTFQEAVSQ 495
Qy 688 KINITCNQVOPFLRTYKNSMQGHGTW-----PCRDRAEALHPTS 729
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 496 SLTLVEPESADIPASIRELIVQPIVTDLYEGDRSELPTESEMPDDEPVIQSIPTVVAPLS 555
Qy 730 -----VHALRPADIQVVALGDSLTAANG-----IGSKRPDDLPTVTOYRGLSYS 774
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 556 SGGREDVDGVIIDSDIRDFL--BELDEBQGNRLRLTIWGTAPDM-DSLRQIRPHT 612
Qy 775 AGGDGSLNVTTLPN-----LLEFNRNLTYG 801
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 613 LKSGSRLVGLTLTGHPAKTLENLNRVLDSGRASPAVVALVERACVVLPEMNAALRGK 672
Qy 802 AVGTGANDTNAFLNQAVPGAKAEDLMSQVOTLMQKMDHRVAFHEDWKITVLIIGSD 861
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 673 ASISTDLAQTALIDLVASGEBA-----FYTSPDSIADDSAT-----LKLVPOLIGS-- 719
Qy 862 LCOVCTSNLYSANFVHLRNALDVLAREVPRVLNVLPFL-----NTIRQV 911
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 720 -----GAPFVDSTVR--ELIEAETANHLDTIKDLNLTAAQOSVSNSLRAV 766
Qy 912 FLGN-----PDKCVOQ-----ASVLCNCVLTIRE 936
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 767 HTLNGAFAMTEMPEITQVMOAETIYIKLLTKHQKASVEGCVLSAATAADLTALRS 826
Qy 937 NSOELAEAFSRAYSRSMRELVSGRYDQEDFSVVLQPFQNI--QLFVLADGLPDT 993
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 827 DS--PRIPSEV-SLVSCLELV-----TTPMDGQSQOQYFESVAEQNLKSVBEQCDV 877

```

```

Qy 994 SFEAPDCIHENQKFSQALAAWTNMLLEPGSKTETLDLRAEMPTCTPTQNEPFLTRPN 1053
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 878 SL-----SNDVLTVYDVSVNLVDLEPLNAVE-----PERPIYDRPQGE----- 917
Qy 1054 SNYTPYIKPAIENWGSDFLCTEWKASNSVTSYHQLRPADIKVVAALGDSLTAAGARP 1113
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 918 -----VHAY-----DVVAASST----- 929
Qy 1114 NSSDLPYSWRGLSMSIGCDGNLEHTTLPLNLLKFPNLYLLGFSTWGTAGALNVAEGA 1173
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 930 -----MSDVVAADTNL----- 940
Qy 1174 RAEDMPAQAQWDLVERKNSPDINL-----EKDMKLVTLFI-GVNDICHYCEPNAHLAT 1226
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 941 ---DVAAHQYDLVSG--NESPANELGLLAFDEPAFELVLDLVESSDILLDHCDNLAKT-- 994
Qy 1227 EYVQHIOQALDILSEELPAPFANVVEVMEASLYQGQGGKCMALAAONNCTCLRHSGSS- 1255
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 995 -----NEAPQDRELVLGLQR--DLHTLKGQ--ARMAGINALGDLGHSIESM 1036
Qy 1286 -----LEKQELKXVWNLOHGISPSYWHQ--YTOREDEPAAVVOF 1323
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1037 LESVADYIMLNDRDMKLEY-----SFVYLHQMILTQRRHVVVTP 1078

```

RESULT 11

115278
 hypothetical protein R155.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15278

C:Geisels, C.; Mamsley, P.; Kramer, J.

submitted to the EMBL data library, May 1997

A:Description: The sequence of C. elegans cosmid R155.

A:Reference number: Z18321

A:Accession: T15278

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-913 <GBI>

A:Cross-references: EMBL:AF003390; NID:G2088866; PTD:G2088867; PIDN:AA54270.1; GSPDB:GN

A:Experimental source: strain Bristol N2; clone R155

C:Genetics:

A:Gene: CESP:R155.4

A:Map position: 3

A:Insertions: 47/1; 68/1; 355/1; 719/1; 794/2; 864/1

Query Match 1.8%; Score 141.5; DB 2; Length 913;
 Best Local Similarity 19.2%; Pred. No. 0.44;
 Matches 191; Conservative 132; Mismatches 335; Indels 339; Gaps 48;

```

Qy 646 NSFPAPDCFHFSKSHSRAASAL-----WNN-----MLEPVQKTRHKEF----- 686
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2 SEFP-----FPLAANNHNSFAESLSITTIWLNICHGASTPLVNPGLQEPHHDENVYRGI 57
Qy 687 --NKINITCNQVOPFLRTYKNSMQGHGTWLPORDRAEALHPTSVALRPADIQVVAL 744
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 58 GNLARTTQDELSTIGQPRSMAR-----LNT 85
Qy 745 GDSLTAANGIGSKRPD-----LPDVTQYRGLSYSAGDG- 779
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 86 GISLQGLAAGSIIPDELISEPLNIGLTTQVARIDPVKLTEWVQCVNLTSEPVSDNL 145
Qy 780 -SLEN--VTLPLNIRENRNLTVAGVGTG--ANDTNAFLNQAVP-----GAAED 826
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 146 KTIENHYLAMSISADKVGKF--GAYVVDGDPYTNLLDKFVSQYVTLDIESTAENLKO 202
Qy 827 LMSQVTL-----MQKMDHRVN----- 845
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 203 PISISEMLKDKITKTKTMFYKIQSAITTSIDIKVAMFQFSGDSALNVIKSLPLSS 262
Qy 846 ---FHEDWKVITVLIIGSDLCYCTDSNLYSANFVDHLR-----NALDVLFRE- 891
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 263 VYSAVSEHEKIGKKTISNLAEF--SKSMSSSSNAEELNKATDLKYLNQMFVYHHQK 320

```

```

QY 892 -----VPRVLVYVFLNPTIMROVFLNPGPDCPCVQASVLCNCTVLRNSOELARLEAF 947
Db 321 HGNKVPKIPELID--GHTDITVFAEN-----SWLNDAYDELCSCEILGSLGV 368
QY 948 SRAVRSNRELVSGRYDQ--EDFSVVLQPFONIQ-----LPVLADGLPDTSPFPAPDC 1000
Db 369 --ATDOAKETIMRIGSLATKLSQSLAISPLAGKVDIKIGSLPAESDK----- 415
QY 1001 IHPQKHSQLARLMTNMLEPLGSKTETDLRLAMPTCTQNEPFLRPRNSNY-- 1058
Db 416 -----NLTLFKLLITDM--ETLTSKVALN--EM--YKTIND--LKKDMSDALQRM 459
QY 1059 -PIKPAIENWSDPELCTEMKASNSVPTS HQLRPADIKVVALDLSLTTAVAGAPNNSSD 1117
Db 460 STIYDGVNNTYDOKLQEVKASKDPVY-----NNLLVNNSSISFLNS 503
QY 1118 LPTSRGLSWSIGGDGNLEHTTLLPNTLKKNPYLGCSTWEGTAGLVAAE----- 1171
Db 504 QPTM-----KKTAEISADALK--NPVULAIANOIPEVINSILNTPPEIKFVIS 548
QY 1172 ----GARBDWPAQAMDVERKNSPDINLEKKDKVLTLPFGVNDLCHYCNPBAHLATE 1227
Db 549 AIXKPEPSTVPLV---ITSVAKILPDI--KKDMKNLQTFVSKKN---SNKTK----- 593
QY 1228 YVOHIQOALDILSEELPRAPVNVVEVMELASLYOGGCKCAMLAAONNCTCLRHSGSL 1287
Db 594 -----ESSVDIL--RELKNA---TVQLSVIGSVARGIFRMEQALGLTNDVSNMK--SEPAVV 643
QY 1288 KOELKKNVNMNLOHGISFSFSTWHTQTOREDRAPVV-----OPFOONTLTPLNER 1335
Db 644 KDEMSKYKLD-----GTDKKNMVLVILGLDELEKLASDLNLFKYSKVPKSNST 690
QY 1336 GDTLTLFFSESDCFHPSDRGHAEMAI-----ALMNNNLBEPVGRKTT 1375
Db 691 NLADPAGIFELAAAYKGVPHNFMAIKASVLTLMANANSNLSAGLLPIILN--LESIGLNF 749
QY 1376 SNNFTHSRAKICKSPESPVL--YTLNRSRL--PDQAEABEVL--YNAVVPYACVGLVWG 1431
Db 750 SSSFSKESKQSLKVLDTSPASLNTINNSRLKSTNAQEBEPGALAKHTIYIAGIPLLI 809
QY 1432 I-----IGTVWRCRBGR-----REDP--PMSLRT 1455
Db 810 IAVCGVWTVWRRRSQRRRAAEAPAHNRDPPESRRS 846

```

RESULT 12
 F69189
 Protoporphyrin IX magnesium chelatase (EC 4.99.1.-) - Methanobacterium thermoautotrophicum
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Nov-1999
 C:Accession: F69189
 R:Smith, D.R.; Doncette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jirawan, N.
 K.; S.; Church, G.M.; Daniels, C.U.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 U. Bacteriol. 179, 7133-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functi
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: F69189
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1708 <MTH>
 A:Cross-references: GB:AE000847; GB:AE000666; NID:g2621756; PIDN:AAB85178.1; PID:g262175
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH673
 C:Superfamily: Methanobacterium thermoautotrophicum cobalamin biosynthesis protein N
 :Keywords: Lyase

Query Match	1.7%	Score 133;	DB 2;	Length 1708;
Best Local Similarity	20.7%;	Pred. No. 5;		
Matches 127;	Conservative 83;	Mismatches 202;	Indels 202;	Gaps 29;

```

0Y ELATILFENKKFOCKT--HOLIEGRDTRDFVWVQPFENVDMPKTSGLDPSFA 650
593 ELATILFENKKFOCKT--HOLIEGRDTRDFVWVQPFENVDMPKTSGLDPSFA 650
Db DUTSILIAKQKYPBNTALHQMIDALYYAAAGKTNLENQFTLAKTYNNRGIEFMSWE 240
181 DUTSILIAKQKYPBNTALHQMIDALYYAAAGKTNLENQFTLAKTYNNRGIEFMSWE 240
0Y PDCEHFSSKSHSRASALMNMNMLEPVQCKTTRHKEFNKINITCENOVQFLR----- 702
651 PDCEHFSSKSHSRASALMNMNMLEPVQCKTTRHKEFNKINITCENOVQFLR----- 702
Db P-----ATVEBASPISSEFFLYRDGQRFTEKDEYFTRYPDPAPKPTAVLSTYGSJTEV 292
241 P-----ATVEBASPISSEFFLYRDGQRFTEKDEYFTRYPDPAPKPTAVLSTYGSJTEV 292
0Y TYKXSMOCHGTWLPFCRBRAPBALHPTSVMHARPDIOVVALGOSLTAAGNIGSKPDLP 762
703 TYKXSMOCHGTWLPFCRBRAPBALHPTSVMHARPDIOVVALGOSLTAAGNIGSKPDLP 762
Db TYADAMQ-----QIIDLAVSRGLANVIPGIG--TWSNVV----- 323
293 TYADAMQ-----QIIDLAVSRGLANVIPGIG--TWSNVV----- 323
0Y DVTTQYRGLSASAGDGSLENVTTLPNILEEF-----NRNLTYAAGTGDANDTAFLN- 816
763 DVTTQYRGLSASAGDGSLENVTTLPNILEEF-----NRNLTYAAGTGDANDTAFLN- 816
Db -----INQTMQNIQIOLCMNQNTYNITAR-GIGANTDLSILIG 364
324 -----INQTMQNIQIOLCMNQNTYNITAR-GIGANTDLSILIG 364
0Y QAVPGAKAEMLDSOVQTLMOCKMDDHVRNFHEDMKVITVLIGSSDLCDYCTDSNLYSAAN 876
817 QAVPGAKAEMLDSOVQTLMOCKMDDHVRNFHEDMKVITVLIGSSDLCDYCTDSNLYSAAN 876
Db TSVSAN-----VVEYEQIL-----DGNGLRSUKISTV-----QPVVYSA-- 400
365 TSVSAN-----VVEYEQIL-----DGNGLRSUKISTV-----QPVVYSA-- 400
0Y FVHDLRNALDYLHREVPRLVNLVDPLNPTIMRQVPLGPNPDKCPVOASVLCNVLTLRE 936
877 FVHDLRNALDYLHREVPRLVNLVDPLNPTIMRQVPLGPNPDKCPVOASVLCNVLTLRE 936
Db LVKFLTDSNVVOYE-----ANPKYFV-KANVLIIDMLTFTTG 437
401 LVKFLTDSNVVOYE-----ANPKYFV-KANVLIIDMLTFTTG 437
0Y NSQELARL-EAFSRAYSRSMRELVSGRYDT-----QEDFSVV-----LOPFQNIQ 982
937 NSQELARL-EAFSRAYSRSMRELVSGRYDT-----QEDFSVV-----LOPFQNIQ 982
Db STSGASTYKFRORSNIPVLRAMITTSYTRIGQWIVSEBGFMSWVSVMYQCAQEMQO-Q 496
438 STSGASTYKFRORSNIPVLRAMITTSYTRIGQWIVSEBGFMSWVSVMYQCAQEMQO-Q 496
0Y LPLVADGLPDTSFFAPDCIPHNOQFHOQLARALMTNMLEPFGSTETLIDLRAMPITCPT 1042
983 LPLVADGLPDTSFFAPDCIPHNOQFHOQLARALMTNMLEPFGSTETLIDLRAMPITCPT 1042
Db IEPLAIGVG-----IGSDPET-GAOWDITVITPE 525
497 IEPLAIGVG-----IGSDPET-GAOWDITVITPE 525
0Y QNEBPF-----LRTPRNSN-----YTPPKPAIENMGSGFLCTEWRKASNSVPTSV 1066
1043 QNEBPF-----LRTPRNSN-----YTPPKPAIENMGSGFLCTEWRKASNSVPTSV 1066
Db RIEKLVSRAPFWILQTMANSDBKVALVYNYTP-PGKQNIQASYL-----NVPEST 575
526 RIEKLVSRAPFWILQTMANSDBKVALVYNYTP-PGKQNIQASYL-----NVPEST 575
0Y HOLRPPADIKVVALGDSLTTAVGARPNNSSDLPTSW--RGLSWSIGDGNLETHTTLPN- 1143
1087 HOLRPPADIKVVALGDSLTTAVGARPNNSSDLPTSW--RGLSWSIGDGNLETHTTLPN- 1143
Db IEI-----LKMKAKAGYS-----VGEIPDADALVEMMIKNGINVANMAPEGELEKANSSNA 627
576 IEI-----LKMKAKAGYS-----VGEIPDADALVEMMIKNGINVANMAPEGELEKANSSNA 627
0Y ILKKFNPYLLGFSF 1157
1144 ILKKFNPYLLGFSF 1157
Db ILMPYEDYLAMENT 641
628 ILMPYEDYLAMENT 641

```

```

RESULT 13
T30884
neural specific DNA binding protein - African clawed frog
C|Species: Xenopus laevis (African clawed frog)
C|Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C|Accession: T30884
R|Marine, J.C.; Bellefroid, E.J.; Samain, H.P.; Martial, J.A.; Pieler, T.
submitted to the EMBL Data Library, December 1995
A|Reference number: Z20918
A|Accession: T30884
A|Status: preliminary; translated from GB/EMBL/DBD
A|Molecule type: mRNA
A|Residues: 1-1361 <MAR>
A|Cross-references: EMBL:U42462; NID:g150837; PID:g150838; PIDN:AAA98467.1
C|Genetics:
A|Note: Xg114

```

	1.7%;	Score 131.5;	DB 2;	Length 1361;
Query Match	19.7%;	Pred. No. 4.3;		
Best Local Similarity	140;	Mismatches 498;	Indels 379;	Gaps 58;
Matches 249;	Conservative			
QY	56	NMPKSHVS-----LKPSDIKF-----VAATGNLEIIPPGCTGDEKQWTERPQVC-----	103	
DB	10	NSPTLSMISAARGSPAEVAHEHLKRGIVGLAPPPTGT-----TPREYCHQMA	59	
QY	104	-----MGVMTYLS-----DIIRFSPSVFPPVCHGTGRVIPHNGAE	139	
DB	60	PLASHASPGYGLLIVQSAAGNTSLHLDYITPMDVSRFSRSPRYPRRL-----SKRRALSTISPLS	117	

QY 140 DLWIOAOLVNMKNKENTOLDFOFDMKLTINFEASNACYLCPAQAONGLAGVD----- 194
 DB 118 DASIDLOTMTIRTSNSL-----VAYINSSSSAASSYGLSLGALSPASPF 165
 QY 195 -----ELMGVLDLYLQGEVPPAFVNLVLSVAVSROYHG--TWLSAPAPPCNCSBETT 246
 DB 166 PIHKPCSLAALAQ--QRSLSSFGHTPLHPSPTFASRQCALISAPAPSNSSAPDS 224
 QY 247 RLKVVNMQSYQEAQMNBL--ASSRYSQESFTVVF-QP----- 282
 DB 225 VLNKVSSESVAVSSTVNQVNIHKRSKYKTEEBEADSVRFQPPDHLTLDKEDDKCKOEPE 284
 QY 283 FFYEETSLSEDEPRLDSTLAWHLNRMMEPAGEDE-----PLSVKH-- 327
 DB 285 HIYTNCHMGCCSEPTODOLVHHINDH--GEKKFPCRWQDCSREKPKAQYML 342
 QY 328 -----GRMKPCSQ-----ESPYSYRNSN-----YLTRL 353
 DB 343 VVHMRRTGCEPKHCTFECEGCFKAYSRLNLTHTLSHTGKFPYCDHEGCKAFSNASDR 402
 QY 354 QKQDCKLEVGEAET-RCPEKDPEDYPTSV-----HRLKPADINVGALGDS 400
 DB 403 AKQONRTHSNEKPYICVPGCTKRYTDPSSLRKHVKTVHGEPAAVTKKHRNDIIQK--PS 460
 QY 401 LTAGNGAGSTPGVNLVLTQYRGLSMVSGDE-----NIGVTTLANTLREFNPFLKGFVS 456
 DB 461 LPKENG-----DNEASAKLSGREHSDVSQDECLQTRITITKDMMHQSPP--GGQS 512
 QY 457 GTCKETSPPNAFLQAVAG-----GRAEDLPVQARLVLMKN-----D 494
 DB 513 SCSESEPPYQNTNNIDSGVDVSLAQSLGDLFLGLEFSPVDSVSMQSGRGPAPPET 572
 QY 495 TRHFOE-----DKIITLFIGNDLCPCNDLVHSPONFTDNGKALDIILAE 544
 DB 573 QRHSAGTGAEREIKDNERFILLIYEPNATQ--NRLPTISANGPDVGVSPSVLIN- 629
 QY 545 VPPAF--VNLVTLVTLVNLRELQEKVVCPRMILSLCVCVLPDNDSTELATLIFN 601
 DB 630 -PRAELISMDVLYMMQNLERR-----DSTSTL----- 657
 QY 602 KKEQKTHOLIESGRYDTRDDFTVVOFPFENVDMPTSE--GLPDNSFPAPDCFFHRSK 659
 DB 658 -----SSAYTSRR--SSGISPPYSRRSSSEFQCGRLNNSSDVSDPSTD 703
 QY 660 SHSRAALMMNMLEPVGQKTRTKFENKINITC-----PNQVPLRLTYKXSMQ 709
 DB 704 ASRRSSASQSHGLPNLNLTPAQHYRLKAKYAATGAPPTPLPNMDRIGLRNKLSLMD 763
 QY 710 GHGTWLP-----CRDRAPSA-LHPTSVALRPADIQVNAALDPSLAGGIGSKP 758
 DB 764 GADPELPFPFQOLPVPRRCSDGGNAGLTPMYPHEI--PGNNSRRAS--DPAVARTGIDDKP 820
 QY 759 DDLPEVTVYQGLSYSAGDGLSENVTTL--PNILREFNRLTGYAVGTGDAN----- 809
 DB 821 --LPRFSRPH-----SMNSMNTLHPPSLSERRNGLCQYTCSDGLHHRVYSPR 867
 QY 810 ----DTNAFL-----NOAVPQAKAEDM--SOVQTLMOQMKDHRVNFHEDKVITVLI 857
 DB 868 PPSISEVNAEMALISCDADVPGGD--DDLMLPDDVQYIRSONREARPCNLQTEYS----- 920
 QY 858 GSGDGLCYCTDSNLSYAA-----NFVDHLNALDVLHREVPV-----LVNLVDFLNP 906
 DB 921 -----SPARNQSTTKSFHNTPEQRPAPGALYSRNPALKECCGQTLANQDNNMPV 972
 QY 907 IMROVFLGNPD--KCPVQOASVLCNVLTLRENSQELARLEAFSRAVRSRRELVSGRY 964
 DB 973 QMNEVSSGTVDSVLPKQOQFALA--GVLAVQOQKONAQOYQSFQAAFMQASHNIMGQOE 1029
 QY 965 DTQEDFSVLIQPF-----FONIGLPLYADGLPPTSFPAPDCI--HPNOKHSOLA 1012
 DB 1030 SVQRNISVNGQRFNYLQOQOQMSQOCI--VSSDPTPOQRYSQOSQMSLSRAMQGGQOIS 1088

QY 1013 RALWTNNLEPLGSKT-----ETLDRAMEPITCPTQ-----NEEFRTPRNSNYTP 1059
 DB 1089 PSC--NNNVERPGVHTTAAPSENTLNHQLAHLVGAHPTQGFANPFSVNDGLHNP--NAYTVQF 1146
 QY 1060 IKPAIE 1065
 DB 1147 OKNGLE 1152
 RESULT 14
 E96590
 hypothetical protein T24C10.3 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C/Accession: E96590
 R/Theologis, A.; Ecker, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Greasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huitza, L.
 Nature 408, 816-820, 2000
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marzials,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, I.
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; MUID:21016719; PMID:11130712
 A/Accession: E96590
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-891 <STO>
 A/Cross-references: GB:AEO05173; NID:g9657517; PIDN:AA00872.1; GSPDB:GN00141
 A/Genetics:
 A:Gene: T24C10.3
 A:Map position: 1
 Query Match 1.7%; Score 130.5; DB 2; Length 891;
 Best Local Similarity 19.4%; Pred. No. 2.4;
 Matches 133; Conservative 92; Mismatches 290; Indels 171; Gaps 27;
 QY 156 LQDFQFDMKLTINFEASNACYLCPAQAONGLAGVDL-----MGVL--DY 202
 DB 169 LNTGERSVAKSVNLMSSTLEVLDPYSDLGADKSDDDLDEWFKTASMGFTLGGSY 228
 QY 203 LQGEVPPAFVNLVLSVAVSROYHGTWLSAPAPPCNCSBETTTL--AKVMQ----- 254
 DB 229 KQERKAVQNTADVSS-----GWSSPLNGTGPFPDPAKDAVGLVARENGDD 279
 QY 255 --MSYQEAQMNBLASSRYSQOE-----SETVFOPEF--YETPSLRS--EDPRL 298
 DB 280 DPNP--NGMGFKVAEAKPEKRDLTNKNESNGMGFGFEPVSKLETNNSFQSSVEKETKM 338
 QY 299 Q-----DSTLANHLMNRMMEPAGEDEPLSVHGHGPMKCPSESRYLPSYRN 346
 DB 339 ENGSISEPPSGDVNSGQTSVAFQPSLEIGNEKEEK--EYOTGK--PKGVLPSEFEDE 393
 QY 347 SNVTLRQKQKDLVREGAETRCPEKDPEDYPTSV--VHRL--KPADINVGALGSLT 402
 DB 394 KSSTSDTLVHEDFVLASDEPVAREKTAESPYSISDLISRLSQVEKKAIVILENSAT 453
 QY 403 AGNGAGSTPGVNLVLTQYRGLSMVSGDENIGVTTLANTLREFNPFLKGFVGTGKET 462
 DB 454 ASNEVNGEDD-----SMFQGPCKPITDSGIAGADDF-----STWFO 493
 QY 463 SPNAFLNQAAGRAEDLPVQARLVLDLKNQTRHIFQ--EMKIIITLFIGNDLCDFC 519
 DB 494 GSPPALKMS-----DYTEVVDPEFDDDS--WEFOGPTQVQKXMSRISGNGLMEY- 540
 QY 520 NDLVHYSPOQFTDNIGKALDIILAEVPPAFVNL--VYTLLEVNLRELQEKVVCPRM 575
 DB 541 --KHSSVEVGN-----OSSVPNGEGELHDKTVIRIENDVDQDLHKLKIETLYHI 589
 QY 576 ILRSL-----CPVLKFDNDSTELATLIFNKKFQKQKTHOLIESGRYDTRDF 623

Db 590 ALVHLEKLEKARDKADSDSEVQKCEIEDLQNLNNDVILISGVNLESLQPGSSGMTELY 649
 QY 624 TVVVOFFENV-----DMPKTS-----GLPDSFPAPDC 653
 Db 650 KALQEKFRRLDEDDLTERLLLSAEKDWKSTIELKNAITTLTKIINLGSLEQOSKASTW 709
 QY 654 FHFS--KSHSRASALMNMMLBPVQKTRHKFEKNKINITCPNOVQPLRTYKNSMOGH 711
 Db 710 FEISSTCAOELRRAASIKWQVINKVDQEEILSKPGKSVALSVEIYRVVKILRASTRLY 769
 QY 712 GTWLPGRDARPSALHPTSVHALRPAD 737
 Db 770 KPMI-----LLAPTSNVLAFLD 787

RESULT 15

T15279
 hypothetical protein R155.3 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans
 C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C/Accession: T15279
 R/Gene: C. J. Mamsley, P., Kramer, J.
 submitted to the EMBL Data Library, May 1997
 A/Description: The sequence of C. elegans cosmid R155.
 A/Reference number: Z18321
 A/Accession: T15279
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1165 <GET>
 A/Cross-references: EMBL:AF003390; NID:G208866; PID:G208870; PIDN:AA54273.1; GSPDB:GN
 A/Experimental source: strain Bristol N2; clone R155
 C/Genetics:
 A/Gene: CESP.R155.3
 A/Map position: 3
 A/Introns: 662/2; 741/1; 813/2; 941/2; 998/1; 1049/3

Query Match 1.7%; Score 130.5; DB 2; Length 1165;
 Best local Similarity 19.5%; Pred. No. 3.8;
 Matches 185; Conservative 140; Mismatches 356; Indels 267; Gaps 44;

QY 396 ALGDSITLPAAGNGASTPGNV--LDVLTQYKGLS--WSVGGDENIGTVTLTANILREFNP 450
 Db 15 AYKDEINLTKGLAVDWGSIWTLIDELSNFLELSRLSMNDEMRLHLIKLKASVTYQTLS 74
 QY 451 LKGFSTVGTGKTSPPN-----APLQAVAGRA--EDLPVQAR---RLVDLMKNDTRIHPQE 501
 Db 75 LVKYNYSQKQETVKSSEVVYSLOKVTSSIALSELPSKASNGDKIIDMLNS-----FEK 129
 QY 502 DWKITP-----LFIGNDLDCFCNDL-----VHYSF 527
 Db 130 STKSLTEDBSPTIGLYELKSLRLKRVSSVALKTVSGFLNGSHDILHLFKDLEDKMIKDAY 189
 QY 528 QNETDNIGKALDILHA-----EVPRAFVNLVTVEIVMLRELYQEKVYCPRMILRLSLC 581
 Db 190 RDSHGMALANLQKTFEASTRAKHOSFVSEAVMLEVSK-----IRSKI 234
 QY 582 PCVLKEDDDSTELATLIEPKKQEKTHQLI--ESGRYDREDFTVVVQPFENVN--MP 637
 Db 235 P--KTSGLADKIATL-----PSKIHETLGPVSVALPTSDK-----DKVDAIP 275
 QY 638 KTSSEGLPDSNFFAPDCFHFSKSHSRASALMNMMLBPVQKTRHKFEKNKINITCPNOV 697
 Db 276 LMAQ-LPAQLKALEKLMNFSSEQLKLPENAKKLGNIIEIVNQ-----DGK-----NOL 321
 QY 698 QPFLRTYKNS--MOGHGTWLPGRD---APSAHPSTVHALRPAD-IQVVA--LGDSTL 748
 Db 322 ENTLKMKFSSTVDKELIYTLILPVKDTLELKONPIHETANIISLDEKVEKFAQDMADFE 381
 QY 749 TAGNGIGSKPDD---LPDVTQYKGLSYSGGDSLENTTLNPIREFRNLTGYAVGT 805
 Db 382 TANNSIRQIESSLIIPIVIT-IRQLRLSNDQGLTIVLNNLPMIKTOLS-SLTSY-----435
 QY 806 GDANDTNALNQAIVPGAEDLMSQVQTLMQKKKDHRVNFHEDMKVITVLLIGSDLCDY 865

Db 436 -----INSYKTKHAE-----TDALKDEKIGLHS--RVIGTATRGI -----470
 QY 866 CTDSNLYSANFVDHRLRNALDVLAHREVPVVLVNLVPLNPTIMROVPLGNPDKCPVQAS 925
 Db 471 ---SNQKLVDFKD-LADIGDLVKSEVEVKQQLND-----502
 QY 926 VLCNCVLTRENSQELARLEAFSRAVRSRMRELVSGRYDTQDFSVLQPFQNIQLPV 985
 Db 503 -----ENVANIKALAGIEGQLKTASGEIVG-----YKSVKPV 535
 QY 986 LADGLPDTSPFAPDCIHP---NOKFHSQALRALMTNMLEPLSGKTETLDRAMPTTCPT 1042
 Db 536 TSTSLPDYSTFPLDAKKVGINLKLPAIISMEKLNVLPIRSSRAT-----RSSF 586
 QY 1043 QNEPFLTRNRSYTPPIKPALENMSDPLCTEMKASNSVPTSVHQLRPADIKVVAALGD 1102
 Db 587 QLDGYLLSLQ-----SLSSISLDFT---KFKSGFDGTSGLDGLGVS-FAKYNN 631
 QY 1103 SLTTAVGARPNNSSDLPTSMRGLSWSIGDGNLEHTTLPNILKKPNPY---LLGFSST 1158
 Db 632 TLTKQKAAALAAQKSGGSS---SASGSGGVFESHSTI--VLIVGCIIVICVILIPAS 685
 QY 1159 TWEGTAGLVVAEGARA-RDMPAQAWDL-----VERMKNSPDINLEKDKLVTLFIVN 1211
 Db 686 SWFYLHQDPAARKCPAIRELPWBEVKPVKKDVSKMKNKPERELEK-----TODESIK 740
 QY 1212 DLCHYENPEAHATBYV-----QHIQALDLISELPRAFPNV 1250
 Db 741 DVPKLLLELDMSAEKRYVDCRYEALQDMKKEICIPLSSELLIKKFFGV 788

Search completed: January 6, 2004, 18:57:26
 Job time : 35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2004, 18:33:20 ; Search time 370 Seconds
(without alignments)
790.635 Million cell updates/sec

Title: US-10-054-691-2
Perfect score: 7766
Sequence: 1 MGRPGIFLELLLLGQGT.....RCRRGRREDPPMSLRVAL 1458

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 733937 seqs, 200641211 residues

Total number of hits satisfying chosen parameters: 733937

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7766	100.0	1458	US-10-054-691-2	Sequence 2, Appl1
2	1392	17.9	270	US-10-094-749-2029	Sequence 2029, Ap
3	613.5	7.9	382	US-10-369-493-7025	Sequence 7025, Ap
4	605.5	7.8	414	US-10-369-493-5575	Sequence 5575, Ap
5	560	7.2	148	US-09-764-891-4078	Sequence 4078, Ap
6	463.5	6.0	981	US-10-369-493-5416	Sequence 5416, Ap
7	406.5	5.2	348	US-10-369-493-6523	Sequence 6523, Ap
8	347	4.5	349	US-10-369-493-6522	Sequence 6522, Ap
9	338	4.4	425	US-10-369-493-5161	Sequence 5161, Ap
10	141.5	1.8	799	US-10-104-047-2335	Sequence 2335, Ap
11	136	1.8	300	US-10-156-761-13312	Sequence 13312, A
12	133	1.7	1708	US-10-369-493-1156	Sequence 1156, Ap
13	129.5	1.7	1633	US-10-359-012-4	Sequence 4, Appl1
14	128.5	1.7	2765	US-09-808-602-84	Sequence 84, Appl
15	128.5	1.7	2765	US-09-800-198-72	Sequence 72, Appl

16	127	1.6	787	15	US-10-128-714-8587	Sequence 8587, Ap
17	124	1.6	2828	10	US-09-905-129-21	Sequence 21, Appl
18	124	1.6	2828	10	US-09-991-630-21	Sequence 21, Appl
19	124	1.6	2828	12	US-10-301-822-49	Sequence 49, Appl
20	124	1.6	2828	12	US-10-332-189-126	Sequence 126, App
21	124	1.6	2828	12	US-10-295-027-58	Sequence 58, Appl
22	124	1.6	2828	12	US-10-295-027-1175	Sequence 1175, Ap
23	124	1.6	2828	15	US-10-176-847-54	Sequence 54, Appl
24	124	1.6	2828	15	US-10-177-293-110	Sequence 110, App
25	123.5	1.6	1230	10	US-09-881-752A-150	Sequence 150, App
26	123	1.6	1335	12	US-10-369-493-11196	Sequence 1196, A
27	122.5	1.6	2764	10	US-09-808-602-80	Sequence 80, Appl
28	122.5	1.6	2764	11	US-09-800-198-68	Sequence 68, Appl
29	121.5	1.6	1276	12	US-10-354-774-66	Sequence 66, Appl
30	121.5	1.6	1276	12	US-10-271-012-66	Sequence 66, Appl
31	121.5	1.6	2597	10	US-09-905-129-2	Sequence 2, Appl1
32	121.5	1.6	2597	10	US-09-905-129-13	Sequence 10, Appl
33	121.5	1.6	2597	10	US-09-905-129-13	Sequence 13, Appl
34	121.5	1.6	2597	10	US-09-991-630-10	Sequence 10, Appl
35	121.5	1.6	2597	10	US-09-991-630-13	Sequence 13, Appl
36	121.5	1.6	2597	10	US-09-991-630-13	Sequence 13, Appl
37	120.5	1.6	2843	10	US-09-987-482-1	Sequence 1, Appl1
38	120	1.5	816	12	US-10-369-493-3409	Sequence 3409, Ap
39	119.5	1.5	1192	9	US-09-758-140-6	Sequence 6, Appl1
40	119.5	1.5	1192	9	US-09-758-140-6	Sequence 71, Appl
41	119.5	1.5	1192	15	US-10-060-036-71	Sequence 28, Appl
42	119.5	1.5	2119	12	US-09-769-744A-28	Sequence 32, Appl
43	119.5	1.5	2843	8	US-08-681-219-32	Sequence 30, Appl
44	119.5	1.5	2843	12	US-10-092-138-30	Sequence 30, Appl
45	119.5	1.5	2843	12	US-09-230-111C-30	

ALIGNMENTS

RESULT 1
US-10-054-691-2
; Sequence 2, Application US/10054691
; Publication No. US20020115846A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Mirand, Maricar
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OR INVENTION: NO. US20020115846A1 Human Lipase and Polynucleotides Encoding It
; FILE REFERENCE: LEX-0303-USA
; CURRENT APPLICATION NUMBER: US/10/054,691
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: US 60/264,049
; PRIOR FILING DATE: 2001-01-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1458
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-054-691-2

Query Match 100.0%; Score 7766; DB 14; Length 1458;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRPGIFLELLLLGQGTPOHHSRKSTLEFGOLMPETLKNSPPCPNPKLGVMPSK 60
Db 1 MGRPGIFLELLLLGQGTPOHHSRKSTLEFGOLMPETLKNSPPCPNPKLGVMPSK 60
QY 61 SVHSLKPSDIKFAAGNLEIPDPGTGDLKODMTERPOQVCMGWMTLSLDIRFSPS 120
Db 61 SVHSLKPSDIKFAAGNLEIPDPGTGDLKODMTERPOQVCMGWMTLSLDIRFSPS 120
QY 121 VPMVCHTKRKVI PHDGAEDLWIOAGELVRNKKENIQDFOFDWKLINVFNSAQCYYC 180
Db 121 VPMVCHTKRKVI PHDGAEDLWIOAGELVRNKKENIQDFOFDWKLINVFNSAQCYYC 180

QY 181 PSAOONGLAAGVDELMGVLDYLOEVPRAVNLVDLSEVAEVSROYHGTWLSPAPEPCN 240
 DB 181 PSAOONGLAAGVDELMGVLDYLOEVPRAVNLVDLSEVAEVSROYHGTWLSPAPEPCN 240
 QY 241 CSEETRLAKVWOMVYOEAANSLAASRYSEGESFTVVVFOPEFYETPSLSHSEDRLOD 300
 DB 241 CSEETRLAKVWOMVYOEAANSLAASRYSEGESFTVVVFOPEFYETPSLSHSEDRLOD 300
 QY 301 STTLAMHLMNRMEPEGEKDEPLSVYHGRPMKCPSESPYLFSYRNSNTLRLQKQODLU 360
 DB 301 STTLAMHLMNRMEPEGEKDEPLSVYHGRPMKCPSESPYLFSYRNSNTLRLQKQODLU 360
 QY 361 EYREGAIEICPDPEADPTVPSVHRLKPADINVIGLADSLTAGNGAGSTPGVNLDTLQ 420
 DB 361 EYREGAIEICPDPEADPTVPSVHRLKPADINVIGLADSLTAGNGAGSTPGVNLDTLQ 420
 QY 421 YRGLSNVSGDENIGTVTTLANILREPNLSLKGFSVGTGKETS PNA FLNQA VAGRAEDL 480
 DB 421 YRGLSNVSGDENIGTVTTLANILREPNLSLKGFSVGTGKETS PNA FLNQA VAGRAEDL 480
 QY 481 PVQARLVLMKNDTRIHFQEDMKITTLFGNDLDCDFCNDLVHYSPOFTDNIIGALDI 540
 DB 481 PVQARLVLMKNDTRIHFQEDMKITTLFGNDLDCDFCNDLVHYSPOFTDNIIGALDI 540
 QY 541 LHAEPRAFNLYTVLEIVNLRELVOEKVYCPRMILRSICPCVLKFDNSTELATLIER 600
 DB 541 LHAEPRAFNLYTVLEIVNLRELVOEKVYCPRMILRSICPCVLKFDNSTELATLIER 600
 QY 601 NKKFOEKTQOLIESGRYDREDEFTVVVQPFENVDMPTSEGLPDNSFPAPDCFHFSSXS 660
 DB 601 NKKFOEKTQOLIESGRYDREDEFTVVVQPFENVDMPTSEGLPDNSFPAPDCFHFSSXS 660
 QY 661 HSRALALMNMNLEPVQOKTRHFKFNKINITCPNOVOPFLTRYKNSMOGHGWLRCRR 720
 DB 661 HSRALALMNMNLEPVQOKTRHFKFNKINITCPNOVOPFLTRYKNSMOGHGWLRCRR 720
 QY 721 APSALHPTSVAHARPADIQVVAALGDSLTPAGNGISGKPDLPVUTTYQYRLSTYSAGDGS 780
 DB 721 APSALHPTSVAHARPADIQVVAALGDSLTPAGNGISGKPDLPVUTTYQYRLSTYSAGDGS 780
 QY 781 LENVTLLPNILREFNRLTGAVGTGDANDTNAFLNQA VEGAEDLMSQVOTLMQMKD 840
 DB 781 LENVTLLPNILREFNRLTGAVGTGDANDTNAFLNQA VEGAEDLMSQVOTLMQMKD 840
 QY 841 DHRVNHEDMKITTVLIGSDSLCDYCTDSNLYSANPVVDLRLAIDLHREVRVNLV 900
 DB 841 DHRVNHEDMKITTVLIGSDSLCDYCTDSNLYSANPVVDLRLAIDLHREVRVNLV 900
 QY 901 DFLNPTIMROVFLGNPKCPVOQASVLCNCVLTILRENSOELARLEAFSRAYSRSMELV 960
 DB 901 DFLNPTIMROVFLGNPKCPVOQASVLCNCVLTILRENSOELARLEAFSRAYSRSMELV 960
 QY 961 SGRYDQEDFSSVYLQEPFONILQPVLAADGLPDTSPFAPDCIHPNOQKHSGLARALMTNML 1020
 DB 961 SGRYDQEDFSSVYLQEPFONILQPVLAADGLPDTSPFAPDCIHPNOQKHSGLARALMTNML 1020
 QY 1021 EPLGSKTETDLDAENPITPTONEPFLRTPRNSNTYTPKPALEWNGSFLCTEWSKN 1080
 DB 1021 EPLGSKTETDLDAENPITPTONEPFLRTPRNSNTYTPKPALEWNGSFLCTEWSKN 1080
 QY 1081 SVFTSVHOLRPADIKVVAALGDSLTTAVGARPNNSSDLPSTWGLSISIGDGLLEHTT 1140
 DB 1081 SVFTSVHOLRPADIKVVAALGDSLTTAVGARPNNSSDLPSTWGLSISIGDGLLEHTT 1140
 QY 1141 LPLILKKFNLYLGFSTSTWEGTAGLVAAEGARAPMDPAQAWDLVERMKNSPDINLEKO 1200
 DB 1141 LPLILKKFNLYLGFSTSTWEGTAGLVAAEGARAPMDPAQAWDLVERMKNSPDINLEKO 1200
 QY 1201 WKLVTLFIGNVDLCHYENPEALATRYVOHIOALDILSEELPRAVNVVEWELASLY 1260
 DB 1201 WKLVTLFIGNVDLCHYENPEALATRYVOHIOALDILSEELPRAVNVVEWELASLY 1260
 QY 1261 QGGGKCMALAAQNNCTCLRHSSSLKQELKKVNNMLQHGISFSYWHQYTORDEPAVY 1320

DB 1261 QGGGKCMALAAQNNCTCLRHSSSLKQELKKVNNMLQHGISFSYWHQYTORDEPAVY 1320
 QY 1321 VOPFFONTITPLNEGDITDLTFSSDCFRFSGRAEWAIALAMNNMLEVGKRTSSNFT 1380
 DB 1321 VOPFFONTITPLNEGDITDLTFSSDCFRFSGRAEWAIALAMNNMLEVGKRTSSNFT 1380
 QY 1381 HSRALKKCPSPFLYTLRNSRLPDQAEAPVLYMAVPAAGVGVVGIIGTVWRC 1440
 DB 1381 HSRALKKCPSPFLYTLRNSRLPDQAEAPVLYMAVPAAGVGVVGIIGTVWRC 1440
 QY 1441 RRGGRREDPMSLRTVAL 1458
 DB 1441 RRGGRREDPMSLRTVAL 1458

RESULT 2
 US-10-094-749-2029
 ; Sequence 2029, Application US/10094749
 ; Publication No. US20030219741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI
 ; APPLICANT: WAKAMATSU, AI
 ; APPLICANT: SATO, HIROYUKI
 ; APPLICANT: ISHII, SHIZUKO
 ; APPLICANT: YAMAMOTO, JUN-ICHI
 ; APPLICANT: ISONO, YUUKO
 ; APPLICANT: HIO, YURI
 ; APPLICANT: OTSUKA, KAORU
 ; APPLICANT: NAGAI, KEIICHI
 ; APPLICANT: IRIE, RYOTARO
 ; APPLICANT: TAMECHIKA, ICHIRO
 ; APPLICANT: SEKI, NAOHICO
 ; APPLICANT: YOSHIKAWA, TSUTOMU
 ; APPLICANT: OTSUKA, MOTOYUKI
 ; APPLICANT: NAGAHARI, KENJI
 ; APPLICANT: MASUHO, YASUHIKO
 ; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
 ; FILE REFERENCE: 084335/0160
 ; CURRENT APPLICATION NUMBER: US/10/094,749
 ; PRIOR FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 60/350,435
 ; PRIOR FILING DATE: 2002-01-24
 ; PRIOR APPLICATION NUMBER: JP 2001-328381
 ; NUMBER OF SEQ ID NOS: 3181
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 2029
 ; LENGTH: 270
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-094-749-2029
 Query Match 17.9%; Score 1392; DB 12; Length 270;
 Best Local Similarity 100.0%; Pred. No. 2,4e-119; Indels 0; Gaps 0;
 Matches 259; Conservative 0; Mismatches 0

QY 491 MKNDTRIHFQEDMKITTLFGNDLDCDFCNDLVHYSPOFTDNIIGALDILHAEPRAVY 550
 DB 1 MKNDTRIHFQEDMKITTLFGNDLDCDFCNDLVHYSPOFTDNIIGALDILHAEPRAVY 60
 QY 551 NLVTVAIEVNLRELVOEKVYCPRMILRSICPCVLKFDNSTELATLIERNKKFOEKTQ 610
 DB 61 NLVTVAIEVNLRELVOEKVYCPRMILRSICPCVLKFDNSTELATLIERNKKFOEKTQ 120
 QY 611 LIESGRYDREDEFTVVVQPFENVDMPTSEGLPDNSFPAPDCFHFSSXSLSAASALYN 670
 DB 121 LIESGRYDREDEFTVVVQPFENVDMPTSEGLPDNSFPAPDCFHFSSXSLSAASALYN 180
 QY 671 NMLEPVQOKTRHFKFNKINITCPNOVOPFLTRYKNSMOGHGWLRCRDAPALHPTSV 730

Db 181 NMLEPVQKTRHKFNKINITCENOVQPLRTYKNSMOGHGWLPCRDAPASALHPTSV 240
 Qy 731 HALRPADIQVVAALGDSLT 749
 Db 241 HALRPADIQVVAALGDSLT 259

RESULT 3

US-10-369-493-7025
 ; Sequence 7025, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2002-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 7025
 ; LENGTH: 382
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-10-369-493-7025

Query Match 7.9%; Score 613.5; DB 12; Length 382;
 Best Local Similarity 38.9%; Pred. No. 36-47;
 Matches 139; Conservative 60; Mismatches 111; Indels 47; Gaps 10;

Qy 376 SDTVPSTVRLKPADINIVGALGDSLTAG-----NGAGSTPGNVLDVLTQYRGLSWS 427
 Db 37 SKVPTNVNSVPRADIKLIMALGDSLTIVGGLFSSQANGAGAE--DVAVALVLYRGIAFG 94
 Qy 428 VGGDENIGTITLANILRENPISLKGKSVGTGKETS-----AFINQAVAGRAEDLPVQ 483
 Db 95 AGDCKTLEEHVITPILIKKNPDVFGYSNGIG--SNVWEIARLNVAMGANAADIPGQ 151
 Qy 484 ARRLVDMKNDTR-IHQEDMKITITLFIGNDLCDFENDLVHSPQFTDNIGKALDIIH 542
 Db 152 AKQLVQLQCHTEVNMKEDMKLNTIFGANDICGKRVEDSPYCAODIQAVQIIY 211
 Qy 543 AEVPRAFVNLVTLVLEIVNRELVOEKKVYCPRMILRSICPVLFKFDNSTELATLIEFNK 602
 Db 212 DNVPRVIVSLTGMLHLEMLAQT-DTGHWFQQR-LHDEGCG-----ESNK 254
 Qy 603 KPE-----KTHQLESGRYDTRDFVTVVQPFENVDMPKTSGLPDNSFPAP 651
 Db 255 NFDADIRQACDYNNKXEKQIENDGFEKNDFTYVVOQPMFOPTLIPMENGKFTOKFPAP 314
 Qy 652 DCFHSSKSHSRASALNMNMLEPVQKTRHKFNKI-NITCPNOVQPLRTYKNS 707
 Db 315 DCFHSGMGAHVSTILMNNIILQPVGSKTSVSNMSPVLOTLLACPDAACPRTKPKNS 371

RESULT 4

US-10-369-493-5575
 ; Sequence 5575, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-28
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 5575
 ; LENGTH: 414
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-10-369-493-5575

Query Match 7.8%; Score 605.5; DB 12; Length 414;
 Best Local Similarity 36.8%; Pred. No. 1.9e-46;
 Matches 141; Conservative 57; Mismatches 118; Indels 67; Gaps 10;

Qy 376 SDTVPSTVRLKPADINIVGALGDSLTAGAGSTGCVNVLDVLTQYRGLSWSVGGDENIG 435
 Db 35 SKVPTNVNSVPRADIKLIMALGDSLTIVGGLFSSQANGAGAE--DVAVALVLYRGIAFG 94
 Qy 436 TVTTLANILRENPISLKGKSVGTGKETS--PNAPINQAVAGRAEDLPVQARRVLDLMDND 494
 Db 95 EHTTVANVLRKPNVTLVAGSKIGISENVWEVSHLNGVPAESKDIIGGNALVNTMAH 154
 Qy 495 TRIHQEDMKITITLFIGNDLDFC-----NDLVHSPQFTDNIGKALDIIHAEVP--- 546
 Db 155 SEINVEDMKLVNIFIGANDICVYCEDPYFNSTALHGNATFEKNIILAAVOILQDNLPRRT 214
 Qy 547 -----RAFYN-----LVTLVLEIVNREL--YOEKVVYCPRMILR 578
 Db 215 NFKKLFBSRISVCKTFSSWREFKNSNIYGRITVSLTGFMFMRLRKIDKKYFCEGLHTE 274
 Qy 579 SLCPVLKEDDNSTELATLIEFNKKEQKTHQ-----LIESGRYDTRDFTV 626
 Db 275 B-CDC-----ESNKQPTDDIDGVCFCGYMAEDIONTGLFDNKDDPTFV 318
 Qy 627 VQPFENV-DMPKTSGLPDNSFPAPDCFFSSKSHSRASALNMNMLEPVQKTRHKF 685
 Db 319 VQPFENVGIDPRPASGVVDMTFAPDCFFHSAIGHGNIGMLMNTIVQPVGKQTSVNL 378
 Qy 686 EN-KINITCPNOVQPLRTYKNS 707
 Db 379 SDPSVGLHCPSTWCPEFPPTKNS 401

RESULT 5

US-09-764-891-4078
 ; Sequence 4078, Application US/09764891
 ; Publication No. US20030077808A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC006
 ; CURRENT APPLICATION NUMBER: US/09/764,891
 ; PRIOR FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 10231
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4078
 ; LENGTH: 148
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-764-891-4078

Query Match 7.2%; Score 560; DB 11; Length 148;
 Best Local Similarity 77.9%; Pred. No. 4.6e-43;
 Matches 113; Conservative 11; Mismatches 17; Indels 4; Gaps 3;

Qy 1013 PALMTMMLPISGKSTETLDRALMPITCTONEPPLRTFRNSYTYPIKPALEMGSDPL 1072
 Db 4 QSLMTMMLPISGKSTETLDRALMPITCTONEPPLRTFRNSYTYPIKPALEMGSDPL 63
 Qy 1073 CTEWKAHSVPTSVQOLRPADIKVVALGDSLTAAAGAPNNSSDPLTSMRG-LWSIGG 1131

Db 64 CTEWKASNSVPTSVHQLRPADIKVVAALGDSLTTAVGARNNSSEPHILEGTLGALLEG 123
 QY 1132 DGNL-ETHHTLPLN--LKKENPYLL 1153
 Db 124 MGNLGDHHTAOSHSEVQPLPMLL 148

RESULT 6
 US-10-369-493-5416
 ; Sequence 5416, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 5416
 ; LENGTH: 981
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-10-369-493-5416

Query Match 6.0%; Score 463.5; DB 12; Length 981;
 Best Local Similarity 29.7%; Pred. No. 1.2e-32;
 Matches 118; Conservative 71; Mismatches 169; Indels 39; Gaps 14;

QY 358 DKLEVRG-ARIRCPDPSDPTVPTSVHRLKPADINIVGALGDSLTAAGAGSTPGNVLD 416
 Db 43 DSEYPMGPFNPNCR--QPHKXSESVHQLHPSQIGVIALGDSVVAQAQKS--SILD 98
 QY 417 VLTQYRLSLSVGDENIGVTTTLANTLREPNBLSKGFSTGKETSPNAFLNQAIVGGR 476
 Db 99 LFDQFPVSVFTGDDVLTNEQATFINIFRKAPRIK---GSSDPVQKFDYDFNNAIPGSF 155
 QY 477 AEDLPVQARRLVDMKNDTRHPEQDKIITLTFGNDLQPCNDLVHVSFONPTDNGK 536
 Db 156 SSELPEDALKLASTLKKLGHETSDTAKFVNIIGHNDLCINNETTFGDETFGKSLHS 215
 QY 537 ALDLHAEPFAFVNLTVLEIVNLRELYOEKV--YCPMILRSLLCPVLKPDNSTEL 594
 Db 216 ALTIQTNVRFVFN---IMPPIVAKHSQAHKSKFC-EFSHKTKTSCIFELNEKEYQ- 270
 QY 595 ATLIEFNKKOEKTHOLIE--SGRYDTRDFTVVVQFFENVDMPTSEGLPDNSFPAPD 652
 Db 271 ---NIKQPEQLNEVEQFNQKYGNSSTFAVVIAAPMLDKSIP-LKQOPNIGLALD 325
 QY 653 CEHSSKSHSAASALNNMIEPVGOKTTRHKFNKINI-----TCPNQVQPIRTYKNS 707
 Db 326 CFHLSPIAHDIAAKQIKWGLKEPFDQKT---ITNQLSVGDRFVCPVECFYATLQNS 381
 QY 708 MQGHGTWLPGRD---RAPSALHFTSVYALRPADIQV 740
 Db 382 ENCE-----PSREMFLLVPSAFASPSGLSMPPVPIPV 414

RESULT 7
 US-10-369-493-6523
 ; Sequence 6523, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 6523
 ; LENGTH: 348
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-10-369-493-6523

Query Match 5.2%; Score 406.5; DB 12; Length 348;
 Best Local Similarity 32.2%; Pred. No. 3.2e-28;
 Matches 117; Conservative 66; Mismatches 133; Indels 47; Gaps 11;

QY 1060 IKPAIENWGSDFLCTEMKASNSV-----PTSVHQLRPADIKVVAALGDSLTTAVGARNPN 1113
 Db 15 VKMALE-----KYLNSIDPHETIPDDVNMKPHIRIYIGAMGDSL--IGSRAE 61
 QY 1114 NSSDLPTSMKGLSWSIGDGNLETHHTLPLNLIK---KENPYLLGFSTSTWEG-TAGLVN 1168
 Db 62 NIVGQRORYGNAAFPGMDPEVDRLTVVNIIFRIIAEKGKRLFGSGTGIDYGENGLNV 121
 QY 1169 AAGCARAPMPAQMVLVERKNSPDINLEKMKVTLTFGVND---LCHYCNPEAHUA 1225
 Db 122 AIGMKSDDLIRAKLSVRKANKKEINLENDKVLSTWLGTDVGTGLGRLEDPVP--V 179
 QY 1226 TEVVOHQIALDLTSEELPFAFVNVVEMLASLYOGGCKAMLAQNNCTCLRSQSS 1285
 Db 180 DEYKSHIEKLLYKKNLPRITVISIYGMPAQLQDAQ-----SLKNGKRA 226
 QY 1286 LEKQELKXNNMLQHGISFSYHQYTOR---EDFAVVQPF-FQNTLPPLNERGDTLT 1341
 Db 227 RYVENQKLD-DLSDGRNVSYDFQNNHNSNDFTVVQPFATVETDYSRDEHGVNPT 285
 QY 1342 FFEEDCFHSDRHAEMALYNNMIEPVGRTKTSNNFHSRAKLKCPSESPIYTLRN 1401
 Db 286 FYASDLFHLKSPGHAVALAKHYWNLPEPVGEKTKRADLGDTPKRIYELNEKNCLIKTVGN 345
 QY 1402 SRL 1404
 Db 346 SKM 348

RESULT 8
 US-10-369-493-6522
 ; Sequence 6522, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 6522
 ; LENGTH: 349
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 US-10-369-493-6522

Query Match 4.5%; Score 347; DB 12; Length 349;


```

;      :      :      :      :      :      :      :      :      :      :
Db      NSMALQKITLQFAHFLFELIDSPSLKTLLEIEDFLVTKMLOEYANDYSRMILETFL-- 646
Qy      589 CVLTIRNSGELARLEFSAAYRSMRELVGSGYD-----TOE---DFSVLQPPF 978
;      :      :      :      :      :      :      :      :      :      :
Db      647 -PVTNESSTEDIALALAKAIAITFWGSLAKISREGNFDAFLTHLNOBOLNFSVY-OLLF 704
;      :      :      :      :      :      :      :      :      :      :
Qy      979 QNTQLPVL-----ADGLPPTSFAPDCIHPNQKFSQULARALMTNMLEPLGSKTE 1028
;      :      :      :      :      :      :      :      :      :      :
Db      705 ENLILNINNLNLAGNSQEAANNLNTDQIMNFN-----LILNMQSETSRKT 752
;      :      :      :      :      :      :      :      :      :      :
Qy      1029 TLDLRAMPITCPTONEPFLRT 1050
;      :      :      :      :      :      :      :      :      :      :
Db      753 VLSRSIVDFLT-----EQFLKT 769

```

```

RESULT 11
US-10-156-761-13312
; Sequence 13312, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMIURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13312
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
; US-10-156-761-13312

```

```

Query Match      1.8%; Score 136; DB 15; Length 300;
Best Local Similarity 27.3%; Pred. No. 0.0022;
Matches 51; Conservativity 31; Mismatches 71; Indels 34; Gaps 9;

```

```

Qy      1083 PTVSHQLRPADIKVVAAGSLTTAVGARPNNSSDLPTSWKGLSWSIGDGNLETHHTLP 1142
;      :      :      :      :      :      :      :      :      :      :
Db      48 PTPVWDRSPASL---AAVGDSITRGPDACA-VLSDCE---VSMATGSDADVDSLA--- 96
;      :      :      :      :      :      :      :      :      :      :
Qy      1143 NILKKFPYLLGFGSTSTWEGTAGLVAAEGARADMPAQMDLYERKMSPDINLEKDWK 1202
;      :      :      :      :      :      :      :      :      :      :
Db      97 --VRLGPF--VGAERGSW-----NYAATGARMAADLEPQMERAVARRP-----E 135
;      :      :      :      :      :      :      :      :      :      :
Qy      1203 LVTLFTGVNDLCHYCNBPFAHLATEVYQHIOALDIIEELPFAFVVEVMELASLYOG 1262
;      :      :      :      :      :      :      :      :      :      :
Db      136 LVTVVMGANDACASASAMTSVA--DFRADLEDAATVROALPKTQVMTSSVPMKRLM--S 193
;      :      :      :      :      :      :      :      :      :      :
Qy      1263 QGKRCAM 1269
;      :      :      :      :      :      :      :      :      :      :
Db      194 QGRTSAL 200

```

```

RESULT 12
US-10-369-493-1156
; Sequence 1156, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.

```

```

; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369, 493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1156
; LENGTH: 1708
; TYPE: PRT
; ORGANISM: Methanobacterium thermoautotrophicum
; US-10-369-493-1156

```

```

Query Match      1.7%; Score 133; DB 12; Length 1708;
Best Local Similarity 20.7%; Pred. No. 0.1;
Matches 127; Conservativity 83; Mismatches 202; Indels 202; Gaps 29;

```

```

Qy      593 ELATLIEFNKKFQDEKT--HQIIEGRTYDREDFTVVVQPFENVDPKTSSEGLPDSNFFA 650
;      :      :      :      :      :      :      :      :      :      :
Db      181 DLTSLIAKOKYKQNTALHQMIDCALYAAAGKTINLENOFKALKMYNNRGILPMSSWE 240
;      :      :      :      :      :      :      :      :      :      :
Qy      651 PDCFHSSKSHSRAASALMNNMLEPVQKTRHKFKFNKINITCPNOVQFLR----- 702
;      :      :      :      :      :      :      :      :      :      :
Db      241 P-----ATVPAPPLSEFLYRDQRTKEDYFTRVPLDPAKPTVAVLASYGSGTEV 292
;      :      :      :      :      :      :      :      :      :      :
Qy      703 TYKSMQHGHTWLPCKDRAPSAHPITSVAHLRADIQVVAALGDSLTAGNGISKDDLP 762
;      :      :      :      :      :      :      :      :      :      :
Db      293 TYADAMQ-----QIDALVSGNLVPIVG--TWSSVY----- 353
;      :      :      :      :      :      :      :      :      :      :
Qy      763 DVTYQRTGLSYSGDGSLENVTLLPNILREF-----NRMLTGYAVGTGDANDTNAFLN- 816
;      :      :      :      :      :      :      :      :      :      :
Db      324 -----ILNQGMQNTIOTLCMPQNTNITAIR-GIGNYDPLTSLIGV 364
;      :      :      :      :      :      :      :      :      :      :
Qy      817 QAVPGAKABDLMSQVOTLMQKMKDHRVNFHEDMKVITVLIGSSDLCDYCTDSNLYSAAN 876
;      :      :      :      :      :      :      :      :      :      :
Db      365 TSVSSAN---VYEQIL-----DNGNLIRSLKISTV-----QPVNVYSA-- 400
;      :      :      :      :      :      :      :      :      :      :
Qy      877 FVDHARNALDYLRVPRVLYNLVDFLNPTRMQVFLGNPDKPCVQOASVLCNCVLTRE 936
;      :      :      :      :      :      :      :      :      :      :
Db      401 LVKFLTDASNVQYE-----ANPEKYPV-KANVITDMLTFTTG 437
;      :      :      :      :      :      :      :      :      :      :
Qy      937 NSQELARL-EAFSRAARSSMRRELVGSGRYDT-----OEDFESV-----LQPFQNIQ 982
;      :      :      :      :      :      :      :      :      :      :
Db      438 STTSGASVTKFFPKRSNIPVLRAMITSTYRTTIGQMTVSEEGFMSVYMQCAQPEHQ-Q 496
;      :      :      :      :      :      :      :      :      :      :
Qy      983 LPVTLADGLPDTSEFPAPDCIHPNQKFSQULARALMTNMLEPLGSKTETLDRAMPITCPT 1042
;      :      :      :      :      :      :      :      :      :      :
Db      497 IEPLAIGVE-----IGSDPET-GAQMWDITVTIPE 525
;      :      :      :      :      :      :      :      :      :      :
Qy      1043 QNEPFP-----LTPRNSN-----YTYPIKAIEWGSDFLCTEKASNSPTSV 1086
;      :      :      :      :      :      :      :      :      :      :
Db      526 RIEKLVSRAEFNMTIRLOTMANSDKVAIVYVNP--PGKNIGISYLV-----NVPEST 575
;      :      :      :      :      :      :      :      :      :      :
Qy      1087 HQLRPADIKVVAAGSLTTAVGARPNNSSDLPTSW--RGLSSIGDGNLETHHTLPN- 1143
;      :      :      :      :      :      :      :      :      :      :
Db      576 IEI-----LKKMKAEGIS-----VGEITQDADALVEMIKNGIIVANNAAPGELEKLANSSNA 627
;      :      :      :      :      :      :      :      :      :      :
Qy      1144 ILKKENPYLLGFST 1157
;      :      :      :      :      :      :      :      :      :      :
Db      628 ILWPEYEDYLAMFNT 641

```

```

RESULT 13
US-10-359-012-4
; Sequence 4, Application US/10359012
; Publication No. US20030232419A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: KOLODKIN, Alex L.
; APPLICANT: TERMAN, Jon R.
; APPLICANT: MAO, Tianyi

```

APPLICANT: PASTERKAMP, Ronald J.
 APPLICANT: YU, Hung-Hsiang
 TITLE OF INVENTION: MOLECULES INTERACTING WITH CASL (MICAL) POLYNUCLEOTIDES, POLYPEPT
 TITLE OF INVENTION: AND METHODS OF USING THE SAME
 FILE REFERENCE: JHU1840-3
 CURRENT APPLICATION NUMBER: US/10/359,012
 CURRENT FILING DATE: 2003-02-04
 PRIOR APPLICATION NUMBER: US 60/388,325
 PRIOR FILING DATE: 2002-06-13
 PRIOR APPLICATION NUMBER: US 60/384,302
 PRIOR FILING DATE: 2002-05-30
 PRIOR APPLICATION NUMBER: US 60/354,178
 PRIOR FILING DATE: 2002-02-04
 NUMBER OF SEQ ID NOS: 40
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 4
 LENGTH: 1633
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-359-012-4

Query Match 1.7%; Score 129.5; DB 12; Length 1633;

Best Local Similarity 19.3%; Pred. No. 0.2; Indels 385; Gaps 49;

Matches 222; Conservative 118; Mismatches 424;

18 OCTPQIHTSPRKSTLEGOIWPETIKNSPPPCNPNK-----LGVMPEKSVHSLSKPSD 69
 430 OCTPPELAEERSLV-RILPOTT-----PENINKPEOYTLDPGTRYPVNLNHCVRPHQ 483
 70 IKFVAIIGNLEIPDPDGTDLK-----QDWTERPOQVCMGVT- 108
 484 VKILYITKELEHYPLERIGSVRSVNLRSKESDIRPSKLLTWQOQTEGQVHNVDLTT 543
 109 -----VLSDIITVFSSV-----PMPVCHTKGVYIPHD 136
 544 SWESGALCAIHRFRPELINPDSLNEDDAVENNOIAFDVAEREFGI.PVTTGKEMASAQ 603
 137 GAEDLWI-----OAOELVR-----NMKENLQDLPQFDWKLINVFNSAQCYL 179
 604 EPKLSMVMWLSKFYELFRCTPIRAPVDSWKNYGENADLS-----L 644
 180 CPQAQNGIAGAAGVDELGVLDYLOQEVPPRAFYNLVD-----LSEVA 221
 645 AKSSISN-----NYLNLTPRKTRPRVDGQGTGENDMNRKRGFTNIDEPS 690
 222 EVSRQYHGTWLSAPBPCNCEETRIKAVVMQMSYQEAANSILAS-----SRVSEQ 273
 691 NFFSSRLG-----SNOECSSKEGQNKV-----KSMANQLAKFEESTRNPISMKOE 739
 274 ESTTVFQPPFFYTT-PSLHSEBPRLODST-----TLAMHNMNMEPAGE 318
 740 RRVSGIGKPVLCSSSGPVHSCCPKEAETPSPPLKQFPVVVYGHVL-BELQVSA 798
 319 KOEPLSVKIGRPMKCPQESPYLFSYRN-----SNVLTRLQKPODKLEVR 364
 799 GSECLG-----RPWRARAKSDLOIGTENFATLSTRPRAALSGVLMRLQOVEKILQKR 854
 365 GAIEICPDQDSDTFTSVHRLKPADINVIAGLSDSLTAGNGAGSTPGNVLVDLTOYRGL 424
 855 AONLAREFTHTKIKEXAAH-----LASWFGHGDFFQNTL-----LSKGL 894
 425 SMS-----VGDENIGTITLANILREFNPISLAKGSVVG- KETSPNAFLN 469
 895 SHHPPSPSPRLDPPPAASSSPSTVDS-ASPAKEXKSSPGHFHSHLRTVHP----- 948
 470 QAVAGGAEADLPVQARLYVDMKNDRIHFQED-----WKITITLFIGNDLDFCNDL 522
 949 QLTVGKVSIGIGAAAEVLVNLNMDHRPKQATSPDESWRKSPPLNLGSSDTCYCKKR 1008
 523 VHSFQNTDNIGKALDILHAEPVRAFNVLTVLEIVNRELYQEKVKVCPRMILSLCP 562
 1009 VY-----VMERLAEAGHFHRECFRCST-CATTLRLAAYTFDCEGKFYCK-----P 1054

583 CVLKFDNDSTELATLIEFNKKQOEK-THQILIESGRYDREDPTVVVQPEFENVDMKTS 641
 1055 HFHCKTNSKQRRARLAKQOREEATWQOEAPRDRTPRESSCAVAAL-----GTLE 1107
 642 GLEPNSFPAPDCPHSSKSHRAASAL-WNNMLE-----PVGKTRHKFEKNINTCPN 695
 1108 GSPPVHSLP-----VLHPLIGMLLDWNDSIPESVHLKAGERISQKSAENGRGRLK 1160
 696 QVOPFLRTYKNSMGHGTWLPCCDRAPASLHPFSVALRPADIQVVAALGDS----- 747
 1161 PVREPL-----LP-----RAGEPLPTQRAQOEKMGTPBAQOGERVPPPKSP 1204
 748 -LTAGNGISKPPDLPDVTYQYKLSYSGAD--GSLVENTLTPN-----TLREERN 797
 1205 LRINAIAIRSLLEPL-----LNSBGKKAMAKQESKITLPTQACTRSFSLKRTNSN 1255
 798 LTGYAVGTGDANDTNAFLNQAAPGAKAEDLMGQVOTLMQMKDHRVNFHEMKVITVLI 857
 1256 KQG-----DQHSFGRNQSSAFSPDPALRTHSLRNPSS-----KVEPAL- 1294
 858 GGSDDLCDYCTDSVLYSAANFVDHLRNALDVHREVRVLVNIYDFINPTIMQVFIGN-- 915
 1295 -----RSPCKSIEEV-----PTLLEKVSLOENF 1318
 916 PD--KCPVOQASVLCNVCVTLRENSQELARLEAFSAVSSMRELVGSGR-----YDQ 968
 1319 PDASKPKPKRISLFS--LRKDKSPESPLQES--RQKQIDRLDGSPPKRYLPEDSAQ 1373
 969 DRSVVLQPF 977
 1374 ALEKLLQPF 1382

RESULT 14

US-09-808-602-84

Sequence 84, Application US/09808602

Patent No. US2002015515A1

GENERAL INFORMATION:

APPLICANT: Vernet, Corine A

APPLICANT: Fernandes, Elma

APPLICANT: Shimkete, Richard A

APPLICANT: Herrman, John L

APPLICANT: Majumder, Kumud

APPLICANT: Mishra, Vidhnu

APPLICANT: Mezes, Peter S

APPLICANT: MacDougall, John

TITLE OF INVENTION: No. US2002015515A1el Proteins and Nucleic Acid Encoding Same

FILE REFERENCE: 15966-697 CIP

CURRENT APPLICATION NUMBER: US/09/808,602

CURRENT FILING DATE: 2001-03-14

PRIOR APPLICATION NUMBER: 09/800,198

PRIOR FILING DATE: 2001-03-05

PRIOR APPLICATION NUMBER: 60/186,596

PRIOR FILING DATE: 2000-03-03

NUMBER OF SEQ ID NOS: 114

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 84

LENGTH: 2765

TYPE: PRT

ORGANISM: Rattus norvegicus

US-09-808-602-84

Query Match 1.7%; Score 128.5; DB 10; Length 2765;

Best Local Similarity 17.9%; Pred. No. 0.66; Indels 565; Gaps 71;

Matches 277; Conservative 194; Mismatches 511;

256 SYQEAANSILASRSYSEGSFTVYVQPPYETTPSHSDPR-----LQDSTLAWH 307
 1112 AITFTWKTDAIGQRYGSDAVSVGFYETCPSLTWEKRTALLQGFELDBSNIGWS 1171
 308 LNMNMEPAGEKDEPLSVRG-----RPMKCPG----- 335
 1172 L-----DKHHTLVNVSGLLKGTEGNQFLTQOPALITTSIMGNRRRSISCPGNDL 1222

```

QY 336 QESPYLFS-----YRNSYLTRLQKPOD-----KLEVRGAEIRCPDK----- 373
DB 1223 AEGNKLAPALAVGIDGSLFVGDFNYIRIRIPSRNVTSLLELRNKEFGKSNBPKHYVL 1282
QY 374 --DP-----SPTVTSVHRLKPADINVIAGLDSITAGNAGSTEGNVLDTLOYRGL 424
DB 1283 AVDPVTGSLYVSDTNSRRIRYRK-----SL-----SGAKDLAGNSEVV----- 1320
QY 425 SWSVGDEINIGYTTLANILREFPNLSKGFSGVGTGKETSPPNAPLNQAVA--GGRADLPV 482
DB 1321 -----AGTEQCCLP--FDEARCGDGGKAVADATL 1346
QY 483 QAARLYDLMKND-----TRIHOEDMKITTLFIGNDL-----CDFCNPLVHSPQ 528
DB 1347 MSRGIAVDCXGKLMYFVADATMIRKVDONGIISTLGSNDLTAARPLSCDSMDVAQVRLE 1406
QY 529 NPTDNIGKALDILHAEPVRAFNVLVYLE--IVNIRELYO--EKKVYC----- 572
DB 1407 WPTDLAVNPM-----NSLYVENNVILRITENHOVSIIAGRPMHCQVPGIDYS 1455
QY 573 --PMILRLSCPCVLKPDNSTELATLIEFNKKFOEKTHQILBSGRY-----DTRE 621
DB 1456 LSKLAHSALESASAJAISHTGVALYTETDEKKNRLROVTTNGELICLLGAASDCCKN 1515
QY 622 DFTVVVQPFEN-----VDMPKTSEGLPDNSFPAPDCPFSSKSHR----- 663
DB 1516 DVNVCICSGDDAATADAILNSPSSLAAPDGTIYIDLGNIRIRAVSKKAPVLAFFQYE 1575
QY 664 AAS--ALMNNMLPEVQOKT--TRHKFNKINITCPNOVOPPLRTYKNSMOGHGTMPL 716
DB 1576 AASPEGELEYVFNADGHQVTSLVTEGVLYNFTYSADNDVTELDINNGSL----- 1627
QY 717 CRDRAPALHPTSVHARPADIQVVALAGSLTAGNGISGKPDLPVYTO-----Y 768
DB 1628 -KLRDSSGNMR--HLLMP--DNQOI--TLVTGTNGGAKA-----VSTONELGLMTY 1673
QY 769 RGLS-----YSAGDGSLENT--TLPNILREFNMLGYAVGTGDAN 809
DB 1674 DGNITGLLATSDETGWTTFYDYHREGRLTVTRPTGVVTSILHEMESISVIDIENSXRDN 1733
QY 810 DTNAFLNQAAPGAKAEDLMSQVQ-----TLMOKMKDHRVNFHE----- 848
DB 1734 DVAVITVTLSSVEASVTVVQOVNRNSYQSGNGLRVAVYANGVGSFSEPHVLAGTLPT 1793
QY 849 -----DMKV-----ITVL-----IGSDLCYCTSNLSAANFV 878
DB 1794 IGRCNISLPENGNLSIEWLRKEQIKGKVTIFGRKLRVHGRNLSIDYDRNIRTEKIYD 1853
QY 879 DHLRNALDVLAHREVP-----VLNVLNPLNPTIRKQVFLGNPDRCPVQAQSVLC 928
DB 1854 DHKFTLRITTYDOVGRFLLPSSGLAAVNSYFFNGRLGLQRGAMSEETDIDKOG-- 1910
QY 929 NCVLTLRENSQELARLEAFSRAY--RSSMRELVSGR-----YTOEDFSVVLQP----- 976
DB 1911 -----RIVGRMPADGKVMGSYSLDKSMVLLDQORQYIFREYSSDRLLHVTMPVARHS 1964
QY 977 -----FQONQLP-----VLADGLPD--TSFFAPDCIHPNQKH-----SOLA 1012
DB 1965 MSTHTSIGYIRNINYPESNASYVIFYSDGRILKTSFLGTG--RQVYFKYKGLSKLS 2020
QY 1013 BALM-----TNMLEPLGSKTETLDLRAEMFITCPONEPFLRPPRS----- 1054
DB 2021 EIVYDSTAVTFGDETTGVLMVNLQSGFSCTIRYKVGRLVDKQIYRSEEGMINARF 2080
QY 1055 NNTY-----PIKPAI-----ENMGSDPLCTEMKASNSVPTSVH 1087
DB 2081 DYTYHDNSFRILASIKPISETPLPVLDLYRDEISGKHEHGK--FGVIYYIDINQIITAAV 2139
QY 1088 QL-----RPADIKY-----VAALGDSLTAAV-----GAPPNNSDLPISWRG 1124
DB 2140 TLSKHEDTHGRIRKEVOYEMFRSLMYMTVOYDSWGRVIRELKLGPYANTTK----- 2191

```

```

QY 1125 LSWSIGDGNLETHHTLPLNLLKKNFNYLGFSTSTWEGTAGLN-----VAAGARARDMPA 1180
DB 2192 YTYVDYDGDQLO-----SVAVNDP-----TWRYSVDLNGNHLNPGNSARLMLPL 2237
QY 1181 QANDVERMKNSEPDINLEKD-----WKLVTLPFIGY 1210
DB 2238 R-YDLRDRITRLDDVOYKIDDDGYLCORGSDFEYNSKGLLTRAANKASGWSVQYRYDGV 2296
QY 1211 NDLCHYCENPEALHAREY--VOHIQOALDILSELPRAFVNV--EVMELASLY----- 1260
DB 2297 SRRASXTNLDGHHLLQYFISDLHH-----PRTIITHYNSNSITSLYLYDLOCH 2344
QY 1261 ----QOQSGKAMLAQNNCTCLRHSSQSLLEKOLKKNVNLQHGISPSFYHQYQRED 1316
DB 2345 LFAMESSGEBEYVVASDNGTLP--AVFSINGLMIKQLOYTAVGEI-----YDSONPD 2395
QY 1317 FAVVOPFPONTITPLNBERGDDLTFSBDCFFSFGHAEMALAMNNMLPEVGRKTS 1376
DB 2396 FQWVI--GFHGYLYDPLTK-----LVHFTORDVDVLA--GR-WTS 2430
QY 1377 NNTHSRAKLKCPSPESPLYTLRNSRLPDQAE--EAPVLYMAV 1420
DB 2431 PDYTMWRNVGKEPAPFN--LYMEKNNPLSNELDLKNVYTVDVKSLV 2475

RESULT 15
US-09-800-198-72
; Sequence 72, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Verne, Corrie AM
; APPLICANT: Fernandes, Blma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Hejmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 72
; LENGTH: 2765
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-800-198-72

Query Match 1.7%, Score 128.5, DB 11; Length 2765;
Best Local Similarity 17.9%; Pred. No. 0.66;
Matches 277; Conservative 194; Mismatches 511; Indels 565; Gaps 71;

QY 256 SYEAMNSLASSRYSEQSFYVFPFFYETTPSHSDPR-----LQDSTLAWH 307
DB 1112 AYTFIDKTDAGQRYGSLDAVAVSGFEYETCPSLILWEKRTALLQGFELPSNLGWS 1171
QY 308 LMRMMEPAGEKDEPLSVKHG-----RPMKCPG----- 335
DB 1172 L-----DKHITLVASGILLKGTGENQPLTOOPALITISMGNGRRRSISCPCNGL 1222
QY 336 QESPYLFS-----YRNSYLTRLQKPOD-----KLEVRGAEIRCPDK----- 373
DB 1223 AEGNKLAPALAVGIDGSLFVGDFNYIRIRIPSRNVTSLLELRNKEFGKSNBPKHYVL 1282
QY 374 --DP-----SPTVTSVHRLKPADINVIAGLDSITAGNAGSTEGNVLDTLOYRGL 424
DB 1283 AVDPVTGSLYVSDTNSRRIRYRK-----SL-----SGAKDLAGNSEVV----- 1320
QY 425 SWSVGDEINIGYTTLANILREFPNLSKGFSGVGTGKETSPPNAPLNQAVA--GGRADLPV 482

```

Db 1321 -----AGTGECLP-----FDEARCGDGKAVADATL 1346
Qy 483 QARBLVDMKND-----TRHFOEDWKITLFIGOND-----CDPCNDLVHNSPQ 528
Db 1347 MSRGIAVDNGMLYVVDATMIRKVDONGIITLLSNDLTAVRPLSCDSMVAQVRL 1406
Qy 529 NFDNIGKALDILHAEPRAVNLVTVLE---IVNIRELYQ-----EKKYC----- 572
Db 1407 WPTDLAVNPM-----NSLYVENNVILRITENHQSIIAGRPNHCOVPGIDYS 1455
Qy 573 -PRMILRSICPVLCRKDDNSTELATLIEFNKKFOEKTHQLESGRY-----DTRE 621
Db 1456 LSKLAHSALLESASALAIASHTGVLITETDEKINRLRQVTNGEICLGAASDCCKX 1515
Qy 622 DFTVVVQPFEN-----VDMPTSEGLPNSFFAPDCFFSSKSHR----- 663
Db 1516 DVNCCICYSGDADAVATDAILNSPSSLAVAPDGTIYIADLGNIRIRAVSKNKPVLNAPNOYE 1575
Qy 664 AAS-----ALMNNMLBPVGOKT---TRHKFNKINITCPOVOPLRTYKNSMOGHGTWLP 716
Db 1576 AASPGOEIYVFPADGHOYTSLVGEYLYNFTYSADNDVTELIDNNGSL----- 1627
Qy 717 CRDRAPSAHPTSVHALRPADIQVVAALGDSLTPANGIGSKPDDLDPVTTQ-----Y 768
Db 1628 -KIRBDSGMPR--HILMP-DNQII-----TLTVGTNGGLKA-----VSTQNLGLMTY 1673
Qy 769 RGJS-----YSAGDGSLENT-----TLPHILREFNRLTGAVGTGDAN 809
Db 1674 DGNITGLATKSDETGTTFFDYDHEGLTNTVTRPTGVVTSLHREMEKSIIVDIENSNDN 1733
Qy 810 DTAFLNOAVPAGAKEDLMSOVO-----TLMOCKMDHVRNPFH----- 848
Db 1734 DTVITNLSVEASYVVOQVNSYOLCNGTLRWYANGVSRHSEPHVLAGTLTP 1793
Qy 849 -----DMKV-----ITVL-----IGSDLCYCTDSNLYSANFV 878
Db 1794 IGRCNISLPWENGLNSIEWRLRKEQIKGKVTIFGRKLRVHGRLNLSDYDRNIRTEKIYD 1853
Qy 879 DHRNALDVLHREVP-----VLVNLVDPLNFTIRQVFLGNDPKCPVQASVLC 928
Db 1854 DHRKFTLRITIDQVGRPFLWPSSGLAAVNVSYFFNGRLAGLORGAMSERTDIDKOG--- 1910
Qy 929 NCVLTRENSQELARLEAFSRAV--RSSMRELVSGR-----YDQEDFSVLOP----- 976
Db 1911 -----RIVSRMPADGKWSYSTLDKSMVLLQSOROYIFREYDSSDLHAVTMPSVARHS 1964
Qy 977 -----PFQNIQLP-----VLADGLPD-----TSFFAPDCIHPNQKH-----SOLA 1012
Db 1965 MSHHTSIGYIRNINYPESNASVIFDYSDGRILKTSFLGTG---RQVYKXGKLSKLS 2020
Qy 1013 RALM-----TNMLEPLGSKTETLDLRAEMPTCTPQNEPLRTPRNS----- 1054
Db 2021 EIVYDSTAVTFGYDETTGVLMVNLQSGFSCIRYRKVGPLVKQIYRFSEEGMINARF 2080
Qy 1055 NNTY-----PIKPAI-----EMWGSDFCTEMKASNSVPTSVH 1087
Db 2081 DVTYHDSNFRIASIKPIVISETLPVLDLYRYDEISGKVEHFGK--FGVLYYDINOIITTAVM 2139
Qy 1088 QL-----RPADIKV-----VAALGDSLTTAV-----GARPNNSDLPTSMRG 1124
Db 2140 TLSKHFDTHRIKEVOYEMFRSLMYMTVOYDSMGRIKRELGLGPYANTTK----- 2191
Qy 1125 LSHSIGDGNLEHTHTLPLNLKKFNPYLLGFSTSTWEGTAGLN---VAAEGARADMPA 1180
Db 2192 YTYDYGDGQLQ-----SVAVNDP-----TWRYSYDLNGNLHLNPGNSARLMP 2237
Qy 1181 QAMDLYERMKNSPDINLEKD-----WKLVTLPFGY 1210
Db 2238 R-YDLRDRITRLDVOYKIDDDGYLCORGSDFEYNSKGLLTRAYNKASGWSVOYRYDGV 2296
Qy 1211 NDLCYCEPEALATEY--VOHIOQALDILSEELPRAFNVV--EVMELASY----- 1260

Db 2297 SRRASYKTNIGHNLQYFYSDLNH-----PTRITHVYHNSSEITSLYYDLOGH 2344
Qy 1261 ----OGQGGKAMLAONNCTCLRHQSLSBKQELKXVMNMLQHGISFSYMHQYQRED 1316
Db 2345 LFMESSEBEEYVVASDNTGTPL--AVFSINGMLTKQLOTAYAGEL-----YDSNPD 2395
Qy 1317 FAVVVOFPFONTLTPLNERGDTDLTFPSDCPHFSDRGHAEMALMNNMLBPVGRKTS 1376
Db 2396 FQWVI-GFHGLYDPLTK-----LVHFTQRDYDLA-----GR-WTS 2430
Qy 1377 NNFTSRALKCQSPSPYLYTLRNSRLLPDOAE--EAPVLYMAV 1420
Db 2431 PVTMMRWVKEBAPFN--LYMFKNNPLSNBLDKNYVTVDVKSMLV 2475

Search completed: January 6, 2004, 19:05:43
Job time : 379 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 6, 2004, 18:59:25 ; Search time 31 Seconds
(without alignments)
1989.975 Million cell updates/sec

Title: US-10-054-691-2

Perfect score: 7766
Sequence: 1 MGRPGIFLELLILGQGT.....RCRGRGRREDPPMSRLTVAL 1458

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 4231058 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patente_AA:*
1: /cgn2_6/ptodata/2/1aa/5A COMB .pep:*
2: /cgn2_6/ptodata/2/1aa/5B COMB .pep:*
3: /cgn2_6/ptodata/2/1aa/6A COMB .pep:*
4: /cgn2_6/ptodata/2/1aa/6B COMB .pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS COMB .pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138	1.8	10182	4 US-09-134-001C-3159	Sequence 3159, App
2	136	1.8	1455	2 US-08-726-012B-2	Sequence 2, App1
3	129.5	1.7	1218	4 US-09-589-567-2	Sequence 2, App1
4	127.5	1.6	1016	3 US-09-180-439-8	Sequence 8, App1
5	123.5	1.6	2636	4 US-09-252-991A-25753	Sequence 25753, App
6	122	1.6	855	4 US-09-107-532A-5646	Sequence 5646, App
7	121.5	1.6	1112	3 US-09-353-585-2	Sequence 2, App1
8	120.5	1.6	2842	1 US-07-741-940-7	Sequence 7, App1
9	120.5	1.6	2842	1 US-08-289-548A-7	Sequence 7, App1
10	120.5	1.6	2842	1 US-08-452-654-7	Sequence 7, App1
11	120.5	1.6	2842	4 US-08-449-731-7	Sequence 7, App1
12	120.5	1.6	2843	1 US-07-741-940-2	Sequence 2, App1
13	120.5	1.6	2843	1 US-08-289-548A-2	Sequence 2, App1
14	120.5	1.6	2843	1 US-08-452-654-2	Sequence 2, App1
15	120.5	1.6	2843	1 US-08-452-655B-2	Sequence 2, App1
16	120.5	1.6	2843	1 US-08-452-655B-7	Sequence 7, App1
17	120.5	1.6	2843	2 US-08-370-235A-2	Sequence 2, App1
18	120.5	1.6	2843	3 US-08-450-582-2	Sequence 2, App1
19	120.5	1.6	2843	3 US-08-450-582-7	Sequence 7, App1
20	120.5	1.6	2843	4 US-08-449-731-2	Sequence 2, App1
21	120.5	1.6	2973	2 US-08-821-355A-7	Sequence 7, App1
22	120.5	1.6	2973	2 US-09-003-687A-7	Sequence 7, App1
23	120.5	1.6	2973	3 US-09-136-605-7	Sequence 7, App1
24	120	1.5	15281	2 US-08-471-119A-2	Sequence 2, App1
25	119.5	1.5	1112	3 US-09-353-585-3	Sequence 3, App1
26	118	1.5	1007	3 US-08-961-083-216	Sequence 216, App
27	118	1.5	1007	4 US-09-536-784-216	Sequence 216, App

28	117.5	1.5	956	4 US-09-134-078-63	Sequence 63, App1
29	117.5	1.5	2509	4 US-09-252-991A-16642	Sequence 16642, A
30	116	1.5	1420	2 US-08-540-804-14	Sequence 14, App1
31	116	1.5	1420	2 US-08-218-265-14	Sequence 14, App1
32	116	1.5	1420	3 US-08-521-872-14	Sequence 14, App1
33	116	1.5	1420	3 US-08-590-399-14	Sequence 14, App1
34	116	1.5	1589	3 US-09-356-952-4	Sequence 4, App1
35	114.5	1.5	1447	4 US-09-376-330-17	Sequence 17, App1
36	114.5	1.5	2154	2 US-08-841-349-4	Sequence 4, App1
37	113.5	1.5	968	3 US-09-180-439-3	Sequence 3, App1
38	113.5	1.5	968	3 US-09-180-439-4	Sequence 4, App1
39	112.5	1.4	1180	3 US-09-224-024-28	Sequence 28, App1
40	112.5	1.4	1180	5 PCT-US94-07902-28	Sequence 28, App1
41	112	1.4	1018	5 PCT-US93-11703-75	Sequence 75, App1
42	111.5	1.4	797	3 US-09-086-912-2	Sequence 2, App1
43	111.5	1.4	797	4 US-09-203-453-2	Sequence 2, App1
44	111.5	1.4	797	4 US-09-500-236-2	Sequence 2, App1
45	111.5	1.4	951	3 US-08-816-346-58	Sequence 58, App1

ALIGNMENTS

```

RESULT 1
US-09-134-001C-3159
; Sequence 3159, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134, 001C
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: US 60/064, 964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055, 779
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3159
; LENGTH: 10182
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3159

Query Match      1.8%; Score 138; DB 4; Length 10182;
Best Local Similarity 18.4%; Pred. No. 0.1;
Matches 265; Conservative 187; Mismatches 546; Indels 440; Gaps 61;

QY 153 KENLDPPQEDWTLINVFSSNASQCLCPASAQNGLAAGVDEL-----MGVLY 202
DB 4346 KNNLDVDTPLKIKKQLDEIDQ-----GTNTDGTQDSDVDVYNDLSAALIEKKVVK 4400
QY 203 LQEEVRAVFNLDLSEVAESRQYGTWLSAPAPPCNCESEETTRAKVVMQSYOEA-- 260
DB 4401 LKRNPTGVQVESVANAQVVDLDMARTSLVDPDTQLOEAGNRLENSINQDTDPDQMT 4460
QY 261 -----WNSLASRYS-DOESFTVVFQPFYETTRLSHSEPRDSTTLVNL--WNRM 312
DB 4461 ODSLNANNYDKLAKARQNLKISKVLGGQPTVAIRQNTDEANNAHQALDTPARSQTLNLR- 4519
QY 313 MEPAKGRDEPLSVKGRPMKCSQESPYLFSTRNSVYVLRQPKQKLEVRBAEIRCPD 372
DB 4520 -----EPYINHINNESHLNNAQDNFKAQVNSAPNNTLE 4554
QY 373 --KDPEDTPTSVHRLKP--ADI-----NVIGA-----LGDSLTAGNG--AGSTP 411
DB 4555 TIKNRKDTLNGSTLSESIADYENKQOENYIDASNKRQDYDAVNAAKGLINQTSF 4614
QY 412 GNVLADVLTO---YRGLSWSGDENIGVTYTLA---NLRFFNPSLKGFSVGTGKETS 463
DB 4615 TMSADVIDQKAEDVKRTKTRALDQNGRLBVAKOQALNHLNTLNDLNDQAGQTLTDITNH-S 4673

```

```

QY 464 PNA-FLNQAAGRAEDLPVQARLVLMKN-DTRHFOEDMKTITLFIGNDLDCFCND 521
DB 4674 PNINSVQAQAE--KANTVNTAMTOLKOTIANVDELF--DGNITN--ADKXKDXANN 4725
QY 522 LVHSPONFTDNIKALDILHAEPRAVNLVLEIVNLRELVOEKVYCCPRMILSLC 581
DB 4726 AVNNAKOLINOSDANOQALDPAEINKYTORVNTTKNDLNGDKLAEAK----- 4773
QY 582 PCVLKFDNDSTELATLIEFNKKFOEKTHQLESRYTREFPVVQPFRE-NVDMKTS 640
DB 4774 -----RDANTTIDGLTYLLEAQNNKAKENV--GKASTKNTITSQLODYNOQNTAMQALR 4825
QY 641 EGIJPD-----NSFAPDCFFSSKSHSRASALMNNMLP-----VGOKT--- 680
DB 4826 NSVVDVANNVANSYINEDNEPKRANYQAVTH---AQTLINAGSNPEMSRUVNQKQAV 4882
QY 681 -TRH-----KFE-----NKINIT-----CENOVOPFLRTYKN 706
DB 4883 NTAHQNLHGQOKLEQAQSSANTEIGNLPNLNTQKAKEKELVNSKQRTVEQOLNQAKS 4942
QY 707 SMOGHFWLPCRDAPRALHPTSYHALRPADIQVVAALGSLTGNGIGSK-PDDLDDVT 765
DB 4943 LDSSWGLKSLVAQOPT-VQKTSYVINEDOPEQ--SAYNDSITWGOTTINKTADPVLDKT 4999
QY 766 TOYRGLSYSAGDGLSNVTTLPVILREFNRNLTVAVGTGDANDTAFIN---QAVPG 821
DB 5000 LV-----DNAISNISTKEALHG-EOKLTAKTEALNALTADLANTPOKKAIXT 5048
QY 822 A-----KAEDMSQVQTLMOQKMDHRVNFHEWVKITVLIGSGDLCDYC 866
DB 5049 AINTAHTRTDVTABOSKANOINSAMHTLRQNISDNESVT----- 5087
QY 867 TDSULYNA-----NPFVHLRNALDVLRVPRVLVLYDFLNTIRQVPLGNPDKCPV 921
DB 5088 NESVYINAEPEKOHAFTEALNNAEIVNEQOATLDANSIN----- 5127
QY 922 QOASVLCNVCVTL-----RENS-QELARLEAFSRAVSSMRELVGSGRYDQ--- 967
DB 5128 OKQOAILITKNAIDGEBOLRAKENADOEINTLQOLDQNSKGLINSQTTTEVASQ 5187
QY 968 -----EDFSVYLQPFQONIQLEVLADG---LEPDSFPAEDCIHPNQKFSQALRALMTNML 1020
DB 5188 LAKAKEINKWME-----QLNHLINGKNQMINSKFTINEDANQOAVSNMIAAS----- 5235
QY 1021 EPLGSKETIDLRAEMPTCTPQNEPFLRTPRNSNYTYPIKPAIEMNGSDPLCTEMKASN 1080
DB 5236 EALKNKSQNPELD-----KVTIE-----QAIN 5257
QY 1081 SVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSDLPYTSWGLSWSIGDGNLEHTTT 1140
DB 5258 NINSAINNLN-GEAKLTAKEDAV-----ASINLUSGL----- 5289
QY 1141 LPNILKKFNPLYLGFSTSTWEGTAGLVAAEGARADMPOAW-----DLV 1186
DB 5290 -----TNEOKTKENOAVGAQTRQOVANKLADAEALDQSMQTLRDV 5331
QY 1187 ER-----MKNSPINEKQKVLTLFIGVNDLCHYCEPNPAHAT--EX 1228
DB 5332 MNQNAHSTSNYFNEEDSTOKNTYDNAIDNGSTYIT--GQH-----NPLINKSTIDQT 5381
QY 1229 VOHIQALDILS--EELPR-----APVNVVWELASLYOGQSGKMLAQN 1274
DB 5382 ISRINTAKNDLHGVKEKQORDKGTANOEGOLGYLNDPOKSGEELVNGSNTRSVEEHIN 5441
QY 1275 NCTCLRHSGSLEKQELKVVNNLQHGISSEFSYHQ--YTQREDFAVAVVOPFQNTLTP 1332
DB 5442 EAKSINNAMQLRKVAEKTNVKQSDYINDSTEHQGYDA-----LOEAMNII 5491
QY 1333 NERGDITL--TFSEDCFFHSRDRGHAEALALM---NNMLPEVGRKRTSNFTHSRA 1384
DB 5492 NEIGNPTLNKSEIEOKLQOULTDAQNALQSHLEAKKNNAITGINKLTALINDAQROKA 5549

```

```

RESULT 2
US-08-726-012B-2
; Sequence 2, Application US/08726012B
; Patent No. 5952190
; GENERAL INFORMATION:
; APPLICANT: Hans Joenje, et al.
; TITLE OF INVENTION: CDNA FOR PANCONI ANEMIA COMPLEMENTATION GROUP A
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klaarquist Sparkman Campbell Leigh & Whinston, LLP
; STREET: One World Trade Center, Suite 1600, 121 S.W. Salmon Street
; CITY: Portland
; STATE: OR
; COUNTRY: USA
; ZIP: 97204-2988
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk, 3.5-inch
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS DOS
; SOFTWARE: Mordperfect 5.1+, ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726, 012B
; FILING DATE: 10/04/96
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Richard J. Polley
; REGISTRATION NUMBER: 28,107
; REFERENCE/DOCKET NUMBER: 3812-45520/RJP/DJE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (503) 226-7391
; TELEFAX: (503) 228-9446
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1455
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: Linear
; US-08-726-012B-2

Query Match 1.8%; Score 136; DB 2; Length 1455;
Best Local Similarity 19.3%; Pred. No. 0.0037;
Matches 225; Conservative 133; Mismatches 393; Indels 408; Gaps 58;

QY 143 IOAOELVRNKNENIQLDFOPDKLINVFPNSAQCYLCPESAQONGLAAGVDELMG---- 198
DB 366 LSAELVGHILQEVLEHQ-EVHWQRLVSFVGLVVCF--PEAQO--LLEDVVARLMAQAFE 420
QY 199 -----VLDYIQOEVPRAPVNLVDLSEVAEVS-RQYHGTWLSRPAPFCNGSEET 245
DB 421 SCOLDSNWTAFVLVVRQALLEGPSAFLSYADWFPASFGSTGRYHG-----CSKKA 469
QY 246 -----TEIAKVNVQWMSQOEAAMNSLLASRYSEQESFTVVPQPFYET--TPSLHSEDPRL 298
DB 470 LVPLFTFLSELV-----FESPRYLQVHILHPLVPSKYSLLTDYISLAKTRL 518
QY 299 QDSTTLAMHLMN-----RMMEPAGEKDEPLS-----VKH-GRPKCPSQESPYL 341
DB 519 AD-----LKVSIEMNGLYEDLASAGDITEPHSQALQDVKAIMEPHGNINIVTVWIASIFR 575
QY 342 FSTRNSNYLTRLOKPODKLEVRREGAEIRCPDKPSTDVPFVSYRL-----KRADI 391
DB 576 RPYVSHFLPALTLTPRLPVPYPSRVAFLBSLRADKIPPSLSTVYCOACSAABEKPED- 634
QY 392 NVIGALGDSLTAAGAGSTPGNVLDVLTQYRGLSWSVCGENGTGVTTLANILREFVPSL 451
DB 635 -----AALGVRAEENIS-----AEEPUGQLTALGEIR----- 661
QY 452 KGFSVGTGKETSPN-----AFINQ--AVAGGRAEDLPVQARLVLMKNQDTRHIF 499
DB 662 -----ASMTDPSQKDVISQAQVAIVISERLRAVLGNHEDSSVEISK-IQLSTINTPLRP 713

```

```

QY 500 QEDWKITLTFIGNDLCDPNDLVHSPONFTNIGKALDILHAEPRAVNLVTVLEIV 559
   : : : : :
Db 714 REHIAVDL-----LTSFCQNLMAAS-----SVAP-----738
QY 560 NURELYOEKKVYPRNITLRLC-----PCVLFKPDNDTELATLIEKKFOEKTQOL----611
   : : : : :
Db 739 -----PERGFWAALFVTRTCGRVLPVAVL-----TRLCQLRHOGPSLSAPHVGLAA 786
   : : : : :
QY 612 --IESGRYDREDFTVVQPFENVMKPSBGLP-----DNSFPADCF- 654
   : : : : :
Db 787 LAVHLG--ESRALPREV-----DVGPRPAGLPVPALFDSILCTKTRSLSEF---CLK 835
   : : : : :
QY 655 -----HFSKSHSRAASALWMNMLEPVGOKTRRHFKFNKINITCPNOVOPFLRT 703
   : : : : :
Db 836 FCTAATSYSLCKTSSQSRDLCLCSLSPGLIK-----KFOPLFRLESEKQPLSE 886
   : : : : :
QY 704 YKNSMOGHGTWLPDRAPASAL-----HPTSVHALRPADIQVVALDLSLTAGNGISKPD 759
   : : : : :
Db 887 DVASLWMPRLHLPBSADMQRALSLMTHTRTREVLKEEDVHL--TYQDWHLLELEIQEAD 944
   : : : : :
QY 760 DLBDVTVQ-----YKGLSYSGAGDSLENT--TLPNITLREPNRLTGAVGTGD 807
   : : : : :
Db 945 ALSDTERODFHQWALHFNHFLPSSASGCGCDLQAACTTILVNALMDFHOSRSRY-----D 999
   : : : : :
QY 808 ANDTNAFLNOAVPGAKA--EDLMSOVQTLMOKKMDHRAVNFHEDWKVITVLIGSDLCQY 865
   : : : : :
Db 1000 HSEN---SDLVFGGTGNEEDITSRLQEMVADLE-----LQOD--LIVPLGHTP--- 1042
   : : : : :
QY 866 CTDSNLSYSAANFV-DHRLNALDVL-----HREV--PRVLVNLVDELNPFTMG 909
   : : : : :
Db 1043 -----SQGHFLFELFRRLQALTSQMSVAAQLQORBELMKRILRL----- 1085
   : : : : :
QY 910 QVPLGNPDKCPVQOAVLCLCVLTLENSQELARLEAFSRAYSNRELVGSKRYTQED 969
   : : : : :
Db 1086 -----PSSVYLGCSS--SFQAEOPITARCQFFPLVNSEKRNFCSHGALTOD- 1129
   : : : : :
QY 970 FSVVLQPFQNIQPLVADGLPD--TSFAPDCIHNOKEHSQLADAL--WTNNLEPLGSK 1026
   : : : : :
Db 1130 ---ITHNFRGLNACLRSRDBSLMVDFLAKC---QTKCPRLTSLNVMWPSLEPV--- 1180
   : : : : :
QY 1027 TETLDLRAEMPICTPQNEBFLRTPRNSNYTYPIKPAIE--NMWSDPLCTE----- 1075
   : : : : :
Db 1181 -----LTCRRRHHCQSPLPRE-----LQKLQEGRGFASPLSEASAPARNPD 1223
   : : : : :
QY 1076 WKASNSVPTSVHOLRPADIK-----VVAALGDSLTAVGARPNN 1114
   : : : : :
Db 1224 WLSAALHFAIQVREENIRKQKLDCEREELLVLEFFSLMGLLSHLTS-----NS 1277
   : : : : :
QY 1115 SSDLPSTW-----RGLSW 1127
   : : : : :
Db 1278 TTDLPRAFHVCAILCELEKRRKISW 1302
   : : : : :

```

RESULT 3
US-09-589-567-2

```

; Sequence 2, Application US/09589567
; Patent No. 6479730
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Shi, Jinru
; TITLE OF INVENTION: Maize DNA Ligase II Orthologue and Uses
; FILE REFERENCE: 1125
; CURRENT APPLICATION NUMBER: US/09/589,567
; CURRENT FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: US 60/145,911
; PRIOR FILING DATE: 1999-07-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FaastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1218
; TYPE: PRT

```

ORGANISM: Zea Mays
US-09-589-567-2

Query Match 1.7%; Score 129.5; DB 4; Length 1218;
Best Local Similarity 18.6%; Pred. No. 0.01;
Matches 237; Conservative 151; Mismatches 418; Indels 471; Gaps 60;

```

QY 10 LELLALLGQGTPOIHNSPKSTLEGQIMPELTKNSPP-----CNBNKLVNMPSSKSVHS 64
   : : : : :
Db 90 MKLITLIGFG-----KNSVFTEDATCTDVHVLGMNLGSETSPY 128
   : : : : :
QY 65 LKPSDIK-----FVAALG-----NLEIPDPGTGLEKOD 94
   : : : : :
Db 129 PQBNFVKLEIMMERGAKAVGFVPTGMVYETKKEGFAYRVKDSLEIHLVPS---EHS 185
   : : : : :
QY 95 WTERPQVCKGMVTVLSDIIRFSPSPVPVCHTGKRVYRPHQAEMLQAOGLVNNKME 154
   : : : : :
Db 186 YNE-----LRDYVKFLHP-----KSVIPTAGVDGKLDSKEAVV--- 219
   : : : : :
QY 155 NQLDPQFDMKLIINVFSAASQCYLCPSAQONGLAAGVDLMGVLDYLOQEVPRAFVNL 214
   : : : : :
Db 220 LQKAF-----AGLVDETANKQEFMAFHRS--- 245
   : : : : :
QY 215 VDLSE-----VAVSRQYHGTWLSAPAPBCNCSSEETTRIAKVWQ-----WSY 257
   : : : : :
Db 246 IDATRCCKDVVNCSSQLDGEDAALLPAISASEQDKLMENTMOETLKELSDFLPWSVS 305
   : : : : :
QY 258 QEAMNSLASKSYSEQESTVTVFQ--PFFYETTPSLHSEDPRLQDSTTLAMHLNRIME 314
   : : : : :
Db 306 QEOITDILMSSBDVVKAAISIFFERERDFEEANVSCNG-----T 345
   : : : : :
QY 315 PGEKXDEPLSVNGRPMKCPSESPLYFSYRNSNYTRL-----QKQP--DKLEVRG 365
   : : : : :
Db 346 PYTEKNH--SSNHGSADVSSQESLFSQKVEYSAKLVNISPMMKRPYTPKVKNGS 403
   : : : : :
QY 366 AEIRCPDKPDSPTVP-----SVARLPRADINVLGALDSLTAGN 405
   : : : : :
Db 404 SMTNPKKKGRILTSSTESGRKOSTITNYFIRATEAFAKRCRPERVTEAHQNVENSQ 463
   : : : : :
QY 406 GAGS--TPGNVLDVLTQYGLSWSVG-----DENIGTVTTLANILREPNLSK 452
   : : : : :
Db 464 LADTEKTOQOIVDOLQI-----VDGSMREYAVYLLEKAGDVAVAVDI---FYSSSE 514
   : : : : :
QY 453 GSVGTXKETSNAFLNQ-----VAGGAEDLPVQARLVOLMKNDTPIHQ 500
   : : : : :
Db 515 GNNV---IEVDKNIMQNAQGETADNCSFTMASGSQATP----- 552
   : : : : :
QY 501 EDWKITLTFIGNDLCDPNDLVHSPONFTNIGKALDILH---AEVPRAFVNLVTVL 556
   : : : : :
Db 553 ---KMSNLAVG-----TSLAHGDSANTSLIEKYLEIERACWLAGQAPRYLHARTF 601
   : : : : :
QY 557 EIVNLRELYOEK---KYVC-----PRMILSLCPCVLKF--DNDSTEL--- 594
   : : : : :
Db 602 DIVE-----KEKGKIKITIAIFCNMFRSLALSPDDVLPVAVVCTNKSIPDHMEELNIG 656
   : : : : :
QY 595 ---ATLIEPNKKFQKTHQILIES-GRVYTRDFTVVVOVPPENV--DMKITSGLPDNSF 648
   : : : : :
Db 657 SLVVTTLLESIGTSRKIHMYKTYGDLRCRQNTLAPRSLISIDVSTJAKL----- 711
   : : : : :
QY 649 PAVDCFHFSKSHSRAASALWMNMLEPVGOKTRRHFKFNKINITCPNOVOPFLRTKNSM 708
   : : : : :
Db 712 -----SAISGSGAG-----RRK-----NLVHLHRS----- 733
   : : : : :
QY 709 OGHGTWLPDRAPASALHPTSVHALRPADIQ--VVAALGDSLTAGNGISKPD-DLPDVT 765
   : : : : :
Db 734 -----CRETEMKFLVTVLRNLRIGAMMKTIIPALAHAVGFDBNCADRAVSLGK 785
   : : : : :
QY 766 TQYRGISYSGAGDSLENT--TLPNITLREPNRLTGAVGCDADDTNAFLNOAVP 820
   : : : : :
Db 786 SOLQSLSTEVS--AEAVNTPNLDLIPSLRE-----GTAFSASSLAM---VP 828
   : : : : :
QY 821 GAKAEDMSQVOT--LMQKKDHRVNFHEDWKVITVLIGSDLCQVCTSNLSYSAANFVD 879
   : : : : :

```

Db 829 GTDIPMLARITNGLTQSLKAFH-----GRSFTCEYKDG---QRAQIHK 870
 QY 880 HLRNALDVLHREVPRVLNVLDFLPTIMROVFLGNBDKCPVOOASVLCNC--VLTREN 937
 Db 871 ALGGSVIRFSRQMKESISRPDLVN--MIKEL-----CKEYSSITLDAEVGINRKK 921
 QY 938 SOELARLEAFSAVYRSSMRELVGSGRYDQEDPSVVLQPFQNIQLPVL-----DG 989
 Db 922 GNEL-----MSFQELSSRERGNKHSISI-----ENIKVDICVFVDIMFHDG 964
 QY 990 -----LPTSPFAPDPCIPNOKFHSQALRALMTNMLEPLGSKETETLDAEMPT---- 1038
 Db 965 ERLDCEPLRQRKRYIHDLFEKPGYF-ELAQOLNIEACEASPDNSSTL---ARWDTFFRN 1020
 QY 1039 TCPQNEPFLRTRNSNYTYPIKPAIENMGSDFLCTEWMKASNSVPSVHQLRPADIKV-- 1096
 Db 1021 ACOSSECEGIMKLTLDVADGYSASAKRCDSW-----LKVKR 1054
 QY 1097 --VAALGDSL-TTAVGA 1110
 Db 1055 DYVGAGDSLVLVPIGA 1071

RESULT 4
 US-09-180-439-8
 ; Sequence 8, Application US/09180439
 ; Patent No. 6225532
 ; GENERAL INFORMATION:
 ; APPLICANT: Dixon, Mark S
 ; APPLICANT: Hatzixanthis, Kostas
 ; APPLICANT: Jones, David A
 ; APPLICANT: Jones, Jonathan DG
 ; TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof
 ; FILE REFERENCE: 620 - 53
 ; CURRENT APPLICATION NUMBER: US/09/180,439
 ; EARLIER FILING DATE: 1998-12-06
 ; EARLIER APPLICATION NUMBER: PCT/GB97/01249
 ; EARLIER FILING DATE: 1997-05-08
 ; EARLIER APPLICATION NUMBER: GB 9609681.3
 ; EARLIER FILING DATE: 1996-05-09
 ; EARLIER APPLICATION NUMBER: GB 9619924.5
 ; EARLIER FILING DATE: 1996-09-24
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 8
 ; LENGTH: 1016
 ; TYPE: PRT
 ; ORGANISM: Lycopersicon esculentum
 ; US-09-180-439-8

Query Match 1.64; Score 127.5; DB 3; Length 1016;
 Best Local Similarity 19.64; Pred. No. 0.011;
 Matches 207; Conservative 134; Mismatches 380; Indels 333; Gaps 55;
 QY 40 TLKNSPPPCNPKLGVNMPKSVSLKPSDIKFVAAGNT-----EIPPD 84
 Db 85 TLVAFPPSSLPLENDLSNNNISGITPPE-----IGNLTNVLVDLNTNQSITPPO 138
 QY 85 PGTGDLKQDWT-----RPQV-CMGVMTVLSDIIRYSPSPVPMVCHTGKRVIP 134
 Db 139 --IGSLAKQIIRIFNNHNGFIPBEIGYLRSLTKSLGINFLSGSIPASIGNM----- 190
 QY 135 HDGAEDMIAQOELVRRMKE-----NLQDDFQ-----DWKLINVFSSNASQCYLC 180
 Db 191 -TNLSFFLVENQLSGFIPEIGYLRSLTKSLDINLSGSI PASLGNL--NNLSPLYL- 246
 QY 181 PSAQONGLAAGVDELWGLV-----DYLOQEVPRAFVNLVDSEVAEVSROYHGT 230
 Db 247 ----YNNQSSGSIPEEIGYLRSLTKSLGINFLSGSIPASIGNLNNLSRLDYNKKSLS 302
 QY 231 WLSPAPPCNCSSETTLAKVVMQWSYQEAAMNSILASSRYSSEDSFTVVFQPPFY--ETT 288
 Db 303 ----IPEIGYLRSLT-----YLDIGENALNGSIPASIGNLNNLSPLYLYNNQLS 348

QY 289 PSLSHEDPRLQDSTTLAMHLMNRMMEPAGEKDEPLSVKIGRPMKCPSESPLYFSYRNS 348
 Db 349 GSIPEEIGYLRSLTYL-----DLGENALNGSI-----PA-----SLGINL 383
 QY 349 YLTRLOKPODKLE-----VRGAEIRCPDKDSDTVPVSRHLK----- 387
 Db 384 NLSRLDLNNKLSGSIPEEIGYLRSLTYLDLGENALNGSIPASIGNLNNLSPLYLYNNQL 443
 QY 388 ----PADIVIGAGSLTAGNCA--GSTR-----GNVLVDLITQV---RGLSMVSGDENG 435
 Db 444 SGSIPEEIGYLSLTYLGNNSNGSIPASIGNLNNLSPLYLYNNQSSGSI--PEEIG 500
 QY 436 TTTTLNIIREFPSSIKG---FSVGTGKETSPPAFINQAVAGRAEDLPQARLYDLMK 492
 Db 501 YLSLTYELFLG--NNSNGSIPASIGNLNNLSRYLYNNQSSGSI--SIPASIGNMNLQ- 554
 QY 493 NDTRIHFQEDWKIITLFIGNDICD----FCN---DLVHSPONFTDNIGAL----- 538
 Db 555 -----TLFLSDNDLIGEPFVFCVLTSLVLYMSRNNLKGVPOCLGNISD 600
 QY 539 -----DLHAEVRAFNVLTYVEIV-----NLREL-----YQEK- 569
 Db 601 LHLMSNSFRGELPSSISNL-TSLKIIDFGNNLEGALPOFFGNISSLQVFMQNNKL 659
 QY 570 -----VYCPRMIL-----RSLCPC---VLKFDNSTE-----LATLI 598
 Db 660 SGTLPNFSIGSLJSLNHLGNELADEIRSLDNCCKQLVDLGDVQDLDPFMWLTLP 719
 QY 599 EFN--KKFOEKTHQLESGRYDREDFTVVVOPFFENVMPKT--SEGPDNSFPAPDCE 654
 Db 720 ELRLVRLTSNKLHGPIRSSGAE-----IMFPLRIIDLSRNAFQDIPSLF----- 766
 QY 655 HFSSKSHSRAASLYMNMLEPVQOKTRKFEKINITEPN---QVOPFLRTY-----K 705
 Db 767 -----EHLGMRTVDKTMEEP---SYESYYDSDVVVTKGLELIVRLSLTYIIDLS 817
 QY 706 NSMOGHGTWLPCKDRAPSLHPTSVHALRPADIQ-----VVAALGD-----SLTAG 751
 Db 818 NKREGH-----IPSVL--GDLAIRLVNVAHALGCIYPSLSGSLSTLESIDLSTN 866
 QY 752 NGIGSKPDDL PVT--TQYRGLS--YSAGDGSLENTYTLNPILEFNNRLTGYAVGTGDA 808
 Db 867 QLSGEIPQOLASLTFLVNLNSHNYLQGCIPQGPQRTFRESNSYEGNDLGRYPVSKGC 926
 QY 809 NDTNAFLNOAVPGAKEDLMSQVQITLMQCKKDDHNVNPFEDMKVITVLG--GSDLC----- 863
 Db 927 KDPVSEKNYTVSALDQOESNSEP-----FNDPMK--AALMGYSGLCIGIS 970
 QY 864 ---DYCTDSNLYSAANFVDHLRVALDVLHREYVR 894
 Db 971 IYIILISTGNLRLARIIELEHKITVQRRKOR 1004

RESULT 5
 US-09-252-991A-25753
 ; Sequence 25753, Application US/09252991A
 ; Patent No. 6531795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 25753
 ; LENGTH: 2636
 ; TYPE: PRT

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25753

Query Match 1.6%; Score 123.5; DB 4; Length 2636;

Best Local Similarity 21.5%; Pred. No. 0.2; Mismatches 264; Indels 215; Gaps 39;

Matches 154; Conservative 83; Mismatches 264; Indels 215; Gaps 39;

814 FLNQAIVGAK-----AEDLMQVOTIMQKKDDHNVFHE---DMKV--I 853
1121 FFEQPFPPNQHNMOSLLKREALKALALQALVEN-HDALKRLEFHEHDTGTHAEHA 1179
854 TVLIGSDICDYCTDSNLVSANFVDHLRNALDVLHREVERVLVNLVDFLNPIMQVFL 913
1180 EATLGA-----LWRAEAVD--RQALSLCESSQSL---DLADGLRLSLV 1223
914 GNDKCPVOQASVLCNCVLTRENSQELAEFSRAYRSMR---ELVG-----S 961
1224 ---DMADGQRLVLLVHLLVVDVSWRL-LEDLQRAYQOSLGEAPRLPGKTSPPKAWA 1279
962 GRVDTQ---EDFVVLQPFQNIQLPVLADGLPDTSPFADCIHP---NQKF---HSQ 1010
1280 GRVSEHARGESMAQLO-FWRE---LLEGAPAEI---PCVHPQALQEPRAISVQSR 1329
1011 LARALWTNMLE--PLGSKTETLD--LRAEMPTICPTONEPFLTPRNSVYTPIKPAIEN 1066
1330 FDRSLTERLLKQAPAYRTQVNDLLTLALRVV-----RMSGASSL-VQLEG 1377
1067 WGSDFCTEKKASNP--TSVQLAPRADIKVVAIGDSLT--AVGARPNNSDLPTS 1121
1378 HGRSEELFADIDLSRTVGWFTS--LPPVRLSPVADIGESIKAIKEQLRAIPD----- 1426
1122 WRGLSMS---IGDGNLEHTTLTPNLIKKNPYLLGFSTWEGTAGLVNAEGARAD 1177
1427 -KGLGGLRLYLAGESSARVLAGLPQARTFN--YLGQPDQDPDMLLDPAESAGAEM 1483
1178 MPQAWDLVERMKNSPDIN-----LEKDWKLVTLFTGVNDLCHYCENPEALATEVYQ 1230
1484 DPEAPLD-----NWLSTNGRVFDGELSDIMSSQMGF-----EDQVRLADDYVA 1529
1231 HIOQALDIISF-----LPPAFVNVVEMLASLYGGGKCA 1268
1530 ELTALVDFCCDSPRHGATSPDFLAGLDQARLDALPVALEEVEDIYPLSPMQG----- 1583
1269 ML-----AAQNNCCCLRHSSOSLEKQELKTNMNL---QHGISFSYHQTQEDRA 1318
1584 MEFHSLYEQASSDYINQMRVDSGLDIPREFRA-WQSALDRHALILSGFAMQGLQOPIQ 1642
1319 VVWQPFQNTLTPYLNBERGDTLTFSEDCFHFSDRGIAEMALMNMNMLEPVGRKTTSN 1378
1643 IYVR-----QRLPFAEDLSQAANRDAALLALA-----AARENG 1678
1379 FTHSRAL-----KCPSPSPYLYTLRN-----SLLPDQAEAEV 1416
1679 FELQAPLRLRLLVKTAEGEHLIYTHHLLDGMGNAQLSEVLSSVAGRSPEQL 1734

RESULT 6
US-09-107-532A-5646

Sequence 5646, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street
CITY: Waltham

STATE: Massachusetts
COUNTRY: USA

ZIP: 02354
COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277
INFORMATION FOR SEQ ID NO: 5646:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...855
SEQUENCE DESCRIPTION: SEQ ID NO: 5646:
US-09-107-532A-5646

Query Match 1.6%; Score 122; DB 4; Length 855;

Best Local Similarity 19.4%; Pred. No. 0.029; Mismatches 237; Indels 230; Gaps 36;

Matches 134; Conservative 88; Mismatches 237; Indels 230; Gaps 36;

357 QDKLEVBGAELRCPKD--PSDVTFTSVRLKPADINVIG---LGSILTRNGAGST 410
118 QEAIVLSNMQKRIKPKELVKGDIYLRAGDVVPADIRLEAHDLVSESLTGESEASE 177
411 PGNVLDVLYOYGLSMVSGDENGIVTTLNLRFNFSLKGFSVGTOKETSPNAPLQ 470
178 KSHA--VLTEER---SIDQKMGFSGLVQ---SSAVGVVETENTTEIGK-INQ 225
471 AVAGRAEDLPVQARSLVDLMKN-----DRIHFOEDMKITTLFIGDNDLC 516
226 ALQSVQOQTPPL-IKINKLNQIRPGILCLSLFLVITFTFRYGMNHIL----- 274
517 DPCNDLVHVSPPNFTDNIGKALDIHAEPVRAVNLVTVLEIVNLRELYOEKKVY--CPR 574
275 -----LSATTALVSVNPEGLPAVLTLISGVVHEMAKKAIIKGLPS 317
575 M-----ILRSLC-----PCVLKRDNDSTELATLIERPKKQOET 608
318 VETLSMTVICSDKGTGLTKNEMTVVDVAKEBACVLSIMKCOELKT-----KEOQKT 371
609 HOLIESGRYDREDFTVVVQPFENVDMKTEBGLPDNSFF----- 649
372 ENL--SG---NPEVALIHRTETANLPLKIKAKIPSSSEYKMTMAHEKCAVIVYG 426
650 ADPCFHSKSHSRASALMNMNMLEPVGQKTR-----HKFENKINITCPNOVQPLRT 703
427 ABEVL-FAKSTLSSAEQEMSQTAAEFAKQGVLFAYKVDKQSELTHETLITQ---LT 482
704 YKNSMGHGTWPCGDRAASALHPTSVYHLRPAIDVOVAALDLSLTAAGSIGSKPPDL 763
483 FAGIA--GLIDPPKESAVKA-----VKECOAGISVKMITDQ-----KDTAK 523
764 VTTQYRGSLYSVA-----GQDGSU-----ENVTTLP-----NLRREFNRLTGY 801
524 ALAEOIGLKHAKVLEGIDLDLMSDEBLIQVPIYDVAFARTTPKHKLIVKALQKN--GE 581

QY 802 AVG-TGD-ANDTNAFLNQA---PGAKAEDJMSQVOTLMQKMDHREHEDMKVITVL 856
 DB 582 IYMTGTDGVNDAPALAKRSVDGIAMGKSEVSQADAMV---LGGD---NHTTAKAV--- 633
 QY 857 IGGSDLCDDYCTDSNLYSANFVDHLRNALDLVHREPRVLNVL---VDFLNPTIMRQ--- 910
 DB 634 -----KEGRRIIDNLQKTIINFLPTALAQGLI 660
 QY 911 ----VFLGNPKCPVQQAAYL-CNCVLT 934
 DB 661 LIALMLNRP-LPLSPVQTLVMVMVTTI 687
 RESULT 7
 US-09-353-585-2
 Sequence 2, Application US/09353585
 Patent No. 6287865
 GENERAL INFORMATION:
 APPLICANT: Dixon, Mark S
 Jones, David A
 TITLE OF INVENTION: Plant pathogen resistance genes and uses
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESS: 8th Floor, 1100 No. 6287865th Glebe Road
 CITY: Arlington
 STATE: Virginia
 COUNTRY: United States of America
 ZIP: 22201-4714
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/353,585
 FILING DATE: 15-Jul-1999
 CLASSIFICATION: C12N 15/29, 15/82, A01H 5/00, A01N 65/00, C12Q 1/68
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/930,277
 FILING DATE: 27-OCT-1997
 APPLICATION NUMBER: PCT/GB96/00785
 FILING DATE: 01-APR-1996
 APPLICATION NUMBER: GB 9506658.5
 FILING DATE: 31-MAR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Ms Mary J Wilson
 REGISTRATION NUMBER: 32,955
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4000
 TELEFAX: (703) 816-4100
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1112 amino acids
 TYPE: amino acid
 STRANDEDNESS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Tomato
 STRAIN: Cf2
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-353-585-2
 Query Match 1.6%; Score 121.5; DB 3; Length 1112;
 Best Local Similarity 18.2%; Pred. No.: 0.055;
 Matches 196; Conservative 139; Mismatches 356; Indels 387; Gaps 48;

QY 53 LGVNMPSKSY-----HSIK-----PEDIKFAAIGNLEIPDPGTDLEK 92
 DB 174 LGINIFSGSIPASVGNLNNLSFLYLNNQSGSIPBEISVLRSLTELDSDNLANSGI--- 231
 QY 93 QDWTERPOOVCMGVTVLSDIIRY---FSPSVMPVCHGKRVIPHDGAE----- 140
 DB 232 -----PASIGNNNLSFLFYGNQSGSIPBEICYL-RSLTYLDLSENALNGSIPAS 282
 QY 141 -----LWIOAQL-----VANMKE----- 154
 DB 283 LGNINLSFLFYGNQSGSIPBEIGYLSLNVLGISENALNGSIPASIGNLNNLSRLNL 342
 QY 155 -NLQDFQEDMKLINFFSNASQCYLCPSAQNGLAGVDEIMGVLDYLOQ----- 205
 DB 343 VNNQSGSIPASIGNL--NNLSMLYL-----YNNQSGSIPASIGNLNNLSMLYLNNQL 395
 QY 206 --EVPAPVNLVDLSEVAEVSROYHGTWLSAPDEPCNSEETRLAKVVMQMSYQEAAMS 263
 DB 396 SGSIIPASIGNLNNLSRLYLNNQSG-----SIPBEIGYLSLTYLDLSNNNSING 445
 QY 264 LLAAS-----RYSRQESFTVVPQPFYETTPSLHSEDPRLQDSTTLAMHLMNM 312
 DB 446 FIPASFGNNSMLAFLFYENQLASVPEBIGYLSLNVLDLSENALNGSIPASIGNLNL 505
 QY 313 MEPAGEKDEPLSVYHGRPMKCPSESPLY-----FSYRNSNYLTRLO 354
 DB 506 SR-----LNLVNNQSGSIPBEIGYLSLNVLDLSENALNGSIPASIGNLNNLSRLN 557
 QY 355 KPQDKLEVEGAEIRCPDPDPTVPTSVHRLKPADINIVAGLDSLTAQNGAGSTPG-- 412
 DB 558 LVNNQL-----SGSIPBEIGYLR--SLNDIGLSEMLN-----GSIPASL 595
 QY 413 ---NVLDVLTQYRG-LSMSVSGDENIGTITLANIRFPRPSLKGTSVGTGKTSPPAFL 468
 DB 596 GNINLSMILYLVNNQSGSI--PEEIGYLSL-TYLSLGNNSLNGL-----IPASFA 644
 QY 469 NQAVAGRAEDLPVQARRVLDLMMKNDTRIHFQEDMKITLFIQ--GNDLCPGNDLVHYS 526
 DB 645 NM-----RNQALILNDNNL-----IGEIFSVGNLTSLEVLYM 678
 QY 527 PON-----FTDNIGKAL-----DILHAEPRA 548
 DB 679 PRNNLKGKVPQCLGNISNLQVLSMSSNSFEGELPSSISNLTSQILDFGRNNLEGALPQC 728
 QY 549 FVNLVTVLEIVNRE-----LYQEKVYCPRAIL-----RSLCPG-----YLK 586
 DB 739 FGN-LSLEVFDMQNNKLSGTLPTNFSIGCSLISLHNGNELEDEIPRSLDNCKKLQVLD 797
 QY 587 FDNSTG-----LATIIEFN--KKFOEKTHQLESGRVYTRDFTVVQPFENVDMPK 638
 DB 798 LGNQNLNDTFPMMLGTLPELRVLRLLTSNKLHGIRSRAB-----IMPDRLIIDLSR 850
 QY 639 T--SEGIPDNSPFAPDCHFSKSHSRASALNNMLLEPVQK----- 679
 DB 851 NAFSODLPSTLF-----EHLKMKRTVKTMBEPESEYDSDVVVTKGLELEI 899
 QY 680 -----TRHAKFENKINITCPNQVO--PLRYTKYKSMQGHGTLPCRDAPSAALHP 727
 DB 900 VRIILSLYTVLIDSSNKFEGHIPSVLGLAIRILANSHNALQGY-----IPSLIGS 950
 QY 728 TSVHALRPADIQVVALGDSLTAGNGIGSKPDLPVLT-QVYGLS--YSAGGDSLENV 784
 DB 951 LSI--LESID-----LSFNQSGSIPQOLASLTFLEFLNLSHNYLQGCIPQGPOR 998
 QY 785 TTLPNILREFFRNLTGYAVGTGDANDTNALQAVGAKAEDJMSQVOTLMQKMDHRE 844
 DB 999 RTFESNSYEGNDLREGTPVSKGCKXPVSEKNTVTALEQDSNSER----- 1045
 QY 845 NFHEDMKVITVLIG-GSDIC-----DYCTDSNLYSANFVDHLRNALDLVHREVP 894
 DB 1046 -FNDFWK--AALMGYSGLCIGISMIYIILSTGNLRLWRLRIEKLHKIIMQRRKQR 1100


```

STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/289,548A
FILING DATE: 12-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107,46943
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2842 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: APC
US-08-289-548A-7

```

Query Match 1.64; Score 120.5; DB 1; Length 2842;

Best Local Similarity 19.84; Pred. No. 0.46; Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

```

34 GOLMPETLKNSPPFCNPNKLGVMNPSKSVLSKPSDIKFAAIGNLBIPDPGTDLEK-92
1920 GQKPKIQKSTFP-----QSSDI-----FDRGAATDEK 1950
QY 93 QDWTBPQVCM--GVMTVLSDIIR-----YFSPSVPM-----123
DB 1951 QNPAIENTPVCFSHNSLSLSDIDQNNKNENPIKETPPDQSGPSKQASGVAPKS 2010
QY 124 -----PVC--HTGKRVIPHDGAEADIMIAQELV-----RNM 152
DB 2011 FHVEDTPVCCSRSSLSLSIDSEDL--LQECISSAMPKKKKPSRLKGDNEKHSRRM 2067
QY 153 -----KENLQDPQ-----FDWRLI-----NVFSPNASQ--CYLCPSAQ 185
DB 2068 GGLIGEDLTLDKDIQRPDEHGLSPSENFDMKALQEGANSIVSLHQAALAAACLSRQ- 2126
QY 186 NGLAAGVDELMGVLDLQOEVPRAFYNLVLSVAEVSROYHGTWLSPAPEPCNSEET 245
DB 2127 ---ASSDSDIILSLKSGISLSP--FHLTPDQEKPKPTSNK-----GPRILKPEKKS 2173
QY 246 TRLAKVVMQSY-----QEAMNSILASRSRSEQSFVVFQPFYETTPSLHSEDPQLD 300
DB 2174 TLETKEISESKIKGGKVKYKSLITGKRVANSBEISQOMQOP-----LQANMPSISR 2225
QY 301 STTLAMHLMRMMEPAGEKDEPLSVKRGKRPMSKPSQSPFLFSRYNSNYLTRLOKPPDKL 360
DB 2226 GRIMI-HIRG--VRNSSSSSTSPVS-KKGPKLKTPAKSP--SEGQRTTSPRGAKSPVKS 2279
QY 361 EV---REGAIRCPCDPKS-----DTVPTSVHRLKPADINVLGALGDSLTAAGAGSTP 411
DB 2280 ELSPVARKQTSIQSGSKAPRSRGSBDSTSPRAQ-QPLSLRIPQSPGNSISPGNGISPP 2338
QY 412 GNVLDVLTVRGLSWSVGDEENIGTVTLTANILREFNPSIKGF-SVGTGK--ETSPNAFL 468
DB 2339 NK-----LSQLPRTSSPASTYKSSGSGKSYTSPPGRQ 2372
QY 469 NQAVAGRAEDLPVQARRLVDLMKNDTRHFOED-KWIIPLFIGNDLCPFCNDLVHYSP 527

```

```

DB 2373 SQ-----QNLTKQ-----TGLSKMASSIPRSESSAKGLQNMNGNG-----AAKKVELSR 2417
QY 528 QNFTDINGKALDLILHAVEPPAFVNLTVLEIVLRLRELYQEKYKCYPRMILRSICPCVLKF 587
DB 2418 MSSTKSGSESD--RSERP-----LVVRQSTRIKE-----APSPTR-----2452
QY 588 DDNSTELATLIEFNKKFKQEKTHQLISGRDYDTEDFTVVQVPPFENVDMKTEGLPDNS 647
DB 2453 -----RKLESASFESLSPSSRPASPTRSQAQTVLSPSLPD--2489
QY 648 FPADECFHSSKSHSPAASALMNM--LEPV-----GQKTRHKENKINITCPNQVOP 699
DB 2490 -----MSLSTHSSVQAGGMRKLPPNLSPTIEYNDGRAXKHDLA-RSHSSPSRL-P 2539
QY 700 PLATYKSMOGHGTLPCRDRAPSALHPTSVALRPADIVVAALGDSLTAAGISGKPD 759
DB 2540 INRS-----GTWREHSHSSSLPRVSTWRTSSSSILSASSSESK-----AKSE 2586
QY 760 DLPDVTTVYRGLSYVSGDGSLENTVTLPLNLR-----EFN-RNLGYAVGTGDANDTNA 813
DB 2587 DEKHVNS-----IGTKOSKENQVSAKGTWRKIKENEFSPNTSQTIVSSGATN-----2635
QY 814 FLNQAVGAKAEDLMSQVOTLMQKMD 840
DB 2636 -----GASKTLIYQMAPAVKSTED 2655

```

RESULT 10
US-08-452-654-7
Sequence 7, Application US/08452654
Patent No. 5691454

GENERAL INFORMATION:

```

APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSLYN, GEORF
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THILVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESSES:
ADDRESS: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, NW
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,654
FILING DATE: 25-MAY-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/741,940
FILING DATE: 08-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107,035574
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 7:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2842 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: APC
US-08-452-654-7

```

```

Query Match 1.6%; Score 120.5; DB 1; Length 2842;
Best Local Similarity 19.8%; Pred. No. 0.46;
Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

```

```

QY 34 GQLMPELTAKSPPCPNKLGVMNPKSVHSLKPSDIKFAVAIGNLEIPDPGTGLEK-92
DB 1920 GQKPTLQKQSTFP-----QSSKDI-----PDRGAATDEKL 1950
QY 93 QDWTERPQOCM---GVMTLSDIIR-----YSPSVPM-----123
DB 1951 QNPAIENTPVCFSHNSLSLSLSDIDENNKNENEPKETEPDPSQGEPSKPAQAGYAPKS 2010
QY 124 -----PVC---HTGKRVIPHDGAEDLWIOAEIV-----RNM 152
DB 2011 FHVEDTPVCSRRNSLSLSLSDIDEDDL---LQECISAMPKXXXXKPSRLKGDNEKSPRM 2067
QY 153 -----KENLQDLPQ-----FDWKL1---NVFSPNASQ---CYLCPSAQ 185
DB 2068 GGILGEDLTLDLKDIOEPDEHGLSPDENFDMKALQEGANSIVSSIHQAALAAACLSRQ- 2126
QY 186 NGLAGGVBEIMGVLDVLOQEVPRAFVNLVDLSEVAESQHYGTMLSPAPCECNCSEET 245
DB 2127 ---ASDSDBILSKGSLGSP--FHLTPDQEKPTSK-----GPRILKDEKS 2173
QY 246 TRLAKVVMQWSY-----QEAANSLASRYSEQESFTVVPQPFYETTPSLHSEDPRLD 300
DB 2174 TLETKEIESKGIKGGKYKSLITGKAVNSEISGMQKQ-----LQANMPSISR 2225
QY 301 STTLAHLNRMMEPAGEKDEPLSVKGRPMKCPQSQSPYLFYSRNSNYLTRLOKQDKL 360
DB 2226 GRIMI-HIPG--VRNSSSTSPVS--KKGPRLKTPAKSP--SEGQATTSPPRGAKEPVKS 2279
QY 361 EV-----REGAELICPDKPS-----DTPVSYHRLKPADINVIGALGDSITAGAGSTP 411
DB 2280 ELSPVARQTSQIGSSKAPSRSGSRDSTPRPAQ-QPLSRPIOSPGRNNTSPRNGISPP 2338
QY 412 GNVLDVLTQYRGLSWSVGDENIGVTTLANILREFNPSLKGF-SVGTGK--ETSPNAFL 468
DB 2339 NK-----LSQLRRTSSPSTASTKSSGSGMSTSPGRQM 2372
QY 469 NQAVAGRAEDLPVQARRLVDMKNDTRIRFOED-WKITLFTIGNDLCPNDLVHYS 527
DB 2373 SQ-----QNLTKQ-----TGLSKNASSIPRSASAKGLNQMNNG--AAKVELSR 2417
QY 528 QNTDNIIGKADLIHAEVPRAPNLVYVEIVMLRELYQEKYICPMILRSICPCVLK 567
DB 2418 MSSTKSSGSESD--RSERP-----VLVROSTRIKE-----ASPFLR-----2452
QY 588 DNDSTELATLIEFNKKFOEKTQHLIESGRVDTREDFTVVVQPFENVDMKTESEGLPDNS 647
DB 2453 -----RLEBSASFESLSPSRPASPTRSQAQTPVLSPLPD--2489
QY 648 FPAIDCFHSSKSHSRAASALMNM--LEPV-----GQKTHHKEKNKINITCPNOVOP 699
DB 2490 -----MSLSTHSSVQAGMKRLPENLSPTEIYNDGRPAKHDA-RSHSSPSRL-P 2539
QY 700 FLATYKNSQGHCTWLPBCRRAPSAHPTVHALRPADIDIVAAAGDSLTAAGCIGSKPD 759
DB 2540 INNS-----STWKRREHSHSSSLPRVSTWRTGSSSLTASSSSEK-----AKSE 2586
QY 760 DLDPVTTQYRGLSYSAGDGLSENVTTLPILR-----EFN-RNLGYAVAGTGDAINDTNA 813

```

```

DB 2587 DEKIVNS-----ISGTRQSKENQVSAKQWIKENEFSPNTSQTIVSSGATN-----2635
QY 814 FLNQAVPQAKAEDLMSQVOTLMQXMD 840
DB 2636 -----GASKTLTYQMAPAVAKTED 2655

```

```

RESULT 11
US-08-449-731-7
Sequence 7, Application US/08449731
Patent No. 6413727
GENERAL INFORMATION:
APPLICANT: ALBERTSEN, HANS
ANAND, RAKESH
CARLSON, MARY
GRODEN, JOANNA
HEDGE, PHILIP J.
JOSLYN, GEOFF
KINZLER, KENNETH
MARKHAM, ALEXANDER F.
MAKAMURA, YUSUKE
THLIVERIS, ANDREW

```

```

TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, LTD
STREET: 1001 G Street, NW
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,731
FILING DATE: 25-May-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/289,548
FILING DATE: 12-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107.46943
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2842 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: APC
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-449-731-7
Query Match 1.6%; Score 120.5; DB 4; Length 2842;
Best Local Similarity 19.8%; Pred. No. 0.46;
Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;
QY 34 GQLMPELTAKSPPCPNKLGVMNPKSVHSLKPSDIKFAVAIGNLEIPDPGTGLEK-92
DB 1920 GQKPTLQKQSTFP-----QSSKDI-----PDRGAATDEKL 1950

```

```

QY 93 QDWTERRPOQVCM---GWMTVLSDIIR-----YFSPSVPM-----123
D 1951 QNPAIENTPVCFSHNSLSLSDIDQENNNKENEPKETEPPDSQGPSPQASGYAPKS 2010
QY 124 -----PVC---HTGKRVIPHOAGEDLWIOAQLV-----RNM 152
D 2011 FHVEDTPVCFRNSLSLSLSDIDEDL---LQECISSAMPKKKPRKLGDNKHSPPRM 2067
QY 153 -----KENLOLDFQ-----FDKLI-----NVFFSNASQ---CYLCPSAQ 185
D 2068 GGLIGEDLTLDLKDIDRDPSEHGLSPDSENFDMKALQEGANSIVSSLHQAACLSRQ- 2126
QY 186 NGLAAGVDLMDGLVDYLOQEVPRAFVNLVDLSEVAEVSROYHGTWLSAPAEPCNCEET 245
D 2127 ---ASSSDSILSLKGSISLQSP---FHLTPDQEKPTSNK-----GPRILKPEKS 2173
QY 246 TRLAKVVMQMSY-----QEAANSILASSRYSEOSTVVPQPFYETTPSLHSEDRLOD 300
D 2174 TLETKIESSEKGIKGGKVKYKSLITGKVRNSNEISGQMOP-----LQANMPSISR 2225
QY 301 STLIAHLNRMMEPAGEKDEPLSVKGRPMKCPQSQSPYLFSTRNSNYLTRLOKPODKL 360
D 2226 GRMTI-HIPG--VRNSSSTSPVS-KKGPKLTPAKSP--SEQRTATSPRGAKSPVKS 2279
QY 361 EV-----REGAELRCPDKPS-----DTVPTSVHRLKPADINVIGALGDSLTAAGAGSTP 411
D 2280 ELSPVARQTSQIGSSSKAPRSRSGRSTSPRPAQ-QPLSRPIQSPGRNLSIPGRNGISPP 2338
QY 412 GNVLDVLTQYRGLSWSVGDENIGITVTLTANILREFNSLKGK-SVGTGK--ETSNAFL 468
D 2339 NK-----LSQLPRTSPSTASTSSSGSGMSTYSPGRM 2372
QY 469 NQAVAGRADLPPVQARLVDMKNDRIHFOE-WKITLFLIGNDLCPDNLVHYS 527
D 2373 SQ-----ONLTQKQ-----TGLSKNASSIPRSESASGLQOMNNGN---ANKVELSR 2417
QY 528 QNFEDNIGKALDILHAEPRAFVNLVTVLEIVNRELQYQEKVCPMILRSLCPVLKF 587
D 2418 MSSTKSGSESD-RSERP-----VLRQSTFIKE-----AFSPTR-----2452
QY 588 DDNSTELATLIEFNKKQOEKTHQLESGRYDREDFTVVOPEFENVDMKTESGLPDS 647
D 2453 -----RKLSEASPELSPPSRPASPTPSAQOTVLSPLSD- 2489
QY 648 FPAPDCHESSKSHRASAALNNM---LEPV-----GQKTRHKKFENKINITCPNOVP 699
D 2490 -----MSLSTHSSVQAGWRLPNNLSPTIEYNDGRPAKRHDA-RSHSESPRL-P 2539
QY 700 FLRTYKNSMGHGTWLPFCRDAPSAHPSTVHALRPADIOVNAALGDSLTAAGNGISKPD 759
D 2540 INNS-----GTWKREHSHKSSSLPRVSTWRTSGSSSILSSASSSEK-----AKSE 2586
QY 760 DLDPVTTQYRGLSYSAGDSELENTVTLPNILR-----EFN-RNLGYAVGTGDANDTNA 813
D 2587 DEKHVNS-----ISGTQSKENQVSAKGTWRIKENEFSPTNSTGYTSSGATN-----2635
QY 814 FLNQAEPGAEADIMSQVOTLMQMKD 840
D 2636 -----GAESKTLIYQMAPAVSKTED 2655

```

RESULT 12
US-07-741-940-2
Sequence 2, Application US/07741940
Patent No. 5352775

GENERAL INFORMATION:
APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDE, PHILIP J.
APPLICANT: JOSLYN, GEOFF

```

? APPLICANT: KINZLER, KENNETH
? APPLICANT: MARKHAM, ALEXANDER F.
? APPLICANT: NAKAMURA, YUSUKE
? APPLICANT: THIVERIS, ANDREW
? TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
? TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
? NUMBER OF SEQUENCES: 94
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Banner, Birch, McKie & Beckett
? STREET: 1001 G Street, NW
? CITY: Washington
? STATE: D.C.
? COUNTRY: USA
? ZIP: 20001-4598
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/07/741,940
? FILING DATE: 19920109
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Kagan, Sarah A.
? REGISTRATION NUMBER: 32,141
? REFERENCE/DOCKET NUMBER: 1107.035574
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-508-9100
? TELEFAX: 202-508-9299
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2843 amino acids
? TYPE: AMINO ACID
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-07-741-940-2

```

Query Match 1.6%; Score 120.5; DB 1; Length 2843;
Best Local Similarity 19.8%; Pred. No. 0.46; Indels 311; Gaps 47;
Matches 184; Conservative 104; Mismatches 328;

```

QY 34 GQLMPETLNKSPPCPNKLGVMNPKSVSLKPSDIKFAAIGNLEIPPDPGDLK- 92
D 1921 GQPKPLQKSTFP-----QSSKDI-----PDRGAATDEKL 1951
QY 93 QDWTERRPOQVCM---GWMTVLSDIIR-----YFSPSVPM-----123
D 1952 QNPAIENTPVCFSHNSLSLSDIDQENNNKENEPKETEPPDSQGPSPQASGYAPKS 2011
QY 124 -----PVC---HTGKRVIPHOAGEDLWIOAQLV-----RNM 152
D 2012 FHVEDTPVCFRNSLSLSLSDIDEDL---LQECISSAMPKKKPRKLGDNKHSPPRM 2068
QY 153 -----KENLOLDFQ-----FDKLI-----NVFFSNASQ---CYLCPSAQ 185
D 2069 GGLIGEDLTLDLKDIDRDPSEHGLSPDSENFDMKALQEGANSIVSSLHQAACLSRQ- 2127
QY 186 NGLAAGVDLMDGLVDYLOQEVPRAFVNLVDLSEVAEVSROYHGTWLSAPAEPCNCEET 245
D 2128 ---ASSSDSILSLKGSISLQSP---FHLTPDQEKPTSNK-----GPRILKPEKS 2174
QY 246 TRLAKVVMQMSY-----QEAANSILASSRYSEOSTVVPQPFYETTPSLHSEDRLOD 300
D 2175 TLETKIESSEKGIKGGKVKYKSLITGKVRNSNEISGQMOP-----LQANMPSISR 2226
QY 301 STLIAHLNRMMEPAGEKDEPLSVKGRPMKCPQSQSPYLFSTRNSNYLTRLOKPODKL 360
D 2227 GRMTI-HIPG--VRNSSSTSPVS-KKGPKLTPAKSP--SEQRTATSPRGAKSPVKS 2280
QY 361 EV-----REGAELRCPDKPS-----DTVPTSVHRLKPADINVIGALGDSLTAAGAGSTP 411
D 2281 ELSPVARQTSQIGSSSKAPRSRSGRSTSPRPAQ-QPLSRPIQSPGRNLSIPGRNGISPP 2339

```

QY 412 GNVLDVLTQYRGLSWSVGDENIGVTTLANILREFNPSLKG-
Db 2340 NK-----LSQLPRTSSPSTASIKSSGSGMSTSGRQM 2373
QY 469 NQAVAGRAEDLPVQARRLVDMKNDTRHFOED-WKITLFTGNDLDFCNDLVHYS 527
Db 2374 SQ-----QNLTKQ-----TGLSKNASSIPRESASGKLNQMNNGNG---ANKKVELSR 2418
QY 528 QNFTDNIGKALDILHAEVPRAFVNLVTLVEIVNLRELQYOKKYVCPRMILRSICPCVLK 587
Db 2419 MSSTKSSGSESD--RSERP-----VLVRQSTFIKE-----APSPTR----- 2453
QY 588 DNSTELATLIEFNKKFOEKTQHOLIESGRYDREDFTVVQPPFENVDMPTSEGLPDNS 647
Db 2454 -----RKLSEASPESSISPSRSPASPTRSQAQTPVLSPLPD-- 2490
QY 648 FPAIDCFHSSKSHSRAASALNNM--LEPV-----GQKTRHKEENKINITCPNOVOP 699
Db 2491 -----MSLSTHSSVQAGWRKLPNNLSPTIEYNDGRPAKHDA-RSHSESPERL-P 2540
QY 700 FLRTYKNSQMGHGTWLPCCRBAASALHPTSVAHLRPADIOVVAALGDSLTAAGIGSKPD 759
Db 2541 INNS-----GTWKRSHSKSSSLPRVSTWRRTGSSSILSASESSEK-----AKSE 2587
QY 760 DLPDVTTOYRGLSYSAGDGLSNVTTLPNILR-----EFN-RNLGVAVGTGDANDTNA 813
Db 2588 DEKHVNS-----ISGTQSKENQVSAGTKWIKIKENEFSPNISTGIVSSGATN----- 2636
QY 814 FLNQAIVGAKAEDLMSQVOTLMQMKD 840
Db 2637 -----GAESKTLIYQMAPAVSKTED 2656

RESULT 13
US-08-289-548A-2
Sequence 2, Application US/08289548A
Patent No. 5648212
GENERAL INFORMATION:
APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSIYN, GEOFF
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THLIVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, LTD
STREET: 1001 G Street, NW
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/289,548A
FILING DATE: 12-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107,46943
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2843 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-289-548A-2
Query Match 1.6%; Score 120.5; DB 1; Length 2843;
Best Local Similarity 19.8%; Pred. No. 0.46;
Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;
QY 34 GQMPERTLKNPPPCPNLGVNMPKSHSLKPSDIKVAALGNLEIPDPDTGLEK- 92
Db 1921 GQKPEPLQKOSTP-----QSSKDI-----PDRGAATDEKL 1951
QY 93 QPWTERPQOVCM--GVMTVLSDIIR-----YFSPVPM----- 123
Db 1952 QNPALIENTVCFSSHNSLSLSDIDQENNNKENEPFKETEPDPDQEPKPOASGYAPKS 2011
QY 124 -----PVC--HTGKRVIHPDGAEDLMTQAOELV-----RNK 152
Db 2012 FIVEDTPVCFSSNSSLSSISDSDDL--LOECISSAMPKKKPERLKDNEKHSPPRM 2068
QY 153 -----KENOLDPQ-----FMKLI-----NVFFSASQ--CYLCSAQ 185
Db 2069 GGLIGEDLTLDKDIQRPDSEHGLSPDENFPMKALQEGANSTVSLHQAALAAACLSRQ- 2127
QY 186 NGLAAGVDELGVLDYLOAEVPRAFVNLVDSVAVSRQYHGTWLSAPAPCNCSEET 245
Db 2128 ---ASSDSLSLSKSGISLSP--FHLTPDQEKFTSNK-----GRLTKRGES 2174
QY 246 TRLAKVVMQWY---QEAWSLASSRYSEQSFVVPFPFFETPSLHSEDRLOD 300
Db 2175 TLETKKIESESQIKGKKVYKSLITGVRSNSEISQMKOP-----LQNMPSISR 2226
QY 301 STTLAHLNMMRMEPGEDEPLSVHGRPMKCPSESPLYBYRNSNLTTRQKQODL 360
Db 2227 GRMTI-HIPG--VRNSSSSTSPVS--KKGPPLTPAKSP--SEGQVATTSPPRAKSVVS 2280
QY 361 EV--REGAEIRCPKQDS-----DVTPTVYRLRPADINVIGALGDSLTAAGAGSTP 411
Db 2281 ELSPVARQTSQIGGSKAASRSGSRDSTPSRAQ-QPLSRPIOSPERANSTSPRANISRP 2339
QY 412 GNVLDVLTQYRGLSWSVGDENIGVTTLANILREFNPSLKG-
Db 2340 NK-----LSQLPRTSSPSTASIKSSGSGMSTSGRQM 2373
QY 469 NQAVAGRAEDLPVQARRLVDMKNDTRHFOED-WKITLFTGNDLDFCNDLVHYS 527
Db 2374 SQ-----QNLTKQ-----TGLSKNASSIPRESASGKLNQMNNGNG---ANKKVELSR 2418
QY 528 QNFTDNIGKALDILHAEVPRAFVNLVTLVEIVNLRELQYOKKYVCPRMILRSICPCVLK 587
Db 2419 MSSTKSSGSESD--RSERP-----VLVRQSTFIKE-----APSPTR----- 2453
QY 588 DNSTELATLIEFNKKFOEKTQHOLIESGRYDREDFTVVQPPFENVDMPTSEGLPDNS 647
Db 2454 -----RKLSEASPESSISPSRSPASPTRSQAQTPVLSPLPD-- 2490
QY 648 FPAIDCFHSSKSHSRAASALNNM--LEPV-----GQKTRHKEENKINITCPNOVOP 699
Db 2491 -----MSLSTHSSVQAGWRKLPNNLSPTIEYNDGRPAKHDA-RSHSESPERL-P 2540
QY 700 FLRTYKNSQMGHGTWLPCCRBAASALHPTSVAHLRPADIOVVAALGDSLTAAGIGSKPD 759
Db 2541 INNS-----GTWKRSHSKSSSLPRVSTWRRTGSSSILSASESSEK-----AKSE 2587
QY 760 DLPDVTTOYRGLSYSAGDGLSNVTTLPNILR-----EFN-RNLGVAVGTGDANDTNA 813
Db 2588 DEKHVNS-----ISGTQSKENQVSAGTKWIKIKENEFSPNISTGIVSSGATN----- 2636

QY 814 FLNQAIVGAKAEDLMSQVOTLMQMKD 840
Db 2637 -----GAESKTLIYQMAPAVSKTED 2656

RESULT 14
US-08-452-654-2
; Sequence 2, Application US/08452654
; Patent No. 5691454
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; APPLICANT: ANAND, RAKESH
; APPLICANT: CARLSON, MARY
; APPLICANT: GRODEN, JOANNA
; APPLICANT: HEDGE, PHILIP J.
; APPLICANT: JOSLYN, GEOFF
; APPLICANT: KINZLER, KENNETH
; APPLICANT: MARKHAM, ALEXANDER F.
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: THILVERIS, ANDREW
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4598
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/452,654
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/741,940
; FILING DATE: 08-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107.035574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2843 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-452-654-2

Query Match 1.6%; Score 120.5; DB 1; Length 2843;
Best Local Similarity 19.8%; Pred. No. 0.46;
Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

QY 34 GQMPETLKNSPPCNKNKGVMPSKVSLSKPSDKFPAALGNLEIPDPGTDLEK-92
Db 1921 GQKPKPILOKQSTFP-----QSSKDI-----PDRGATDEKL 1951

QY 93 QDWTERRQOQCM---GVMTVLSDIIR-----YFSPSVPM-----123
Db 1952 QNFAIENTPVCFSHNSLSLSDIDENNKNENPIKETEPDPSQGEPSKPAQSGVAPKS 2011

QY 124 -----PVC---HTGKRVIPHDGAEDLMIQAOELV-----RNM 152
Db 2012 FHVEDPFCVCSRRNSLSLSIDSEDL---LQECISAMPKPKKPRLLKGDNEKHSRRM 2068

QY 153 -----KENTLODPQ-----FDWKLI-----NVEFSNASQ--CYLCPSAQ 185
Db 2069 GGILGEDLTLDKDIDRPSDEHGLSPDSENFWDKALIQEGANSIVSSLHQAAAAACLSRQ- 2127

QY 186 NGIAGGVDELNGVLDYLOQEVPRAFVNLVLSVAEVSRYQYHGTWLSPAPECNCSEET 245
Db 2128 ---ASDSDBIELSLKGSISGSP--FHLRPDQGEKPTSNK-----GPRILKPEKES 2174

QY 246 TRLAKYVMQSY-----QENANSLASRYSSEDSFTVVFPFFETTPSHSEDPRLQD 300
Db 2175 TLETKIESEKIKGCKKYKSLITGKVRNSSEISGQKQP-----LQANMPSISR 2226

QY 301 STLAWHLNRMWEPAGEKDEPLSVKGRPMKCPGSEPLTFGRSNVYLRQLQPODKL 360
Db 2227 GRTMI-HIRG--VRNSSSTSPVS--KKGPPLKTPAKSP--SGQTATTSPRKAKSVKS 2280

QY 361 EV-----REGAETRCPPDKPS-----DTPTSVHRLKPADINVIGALGDSLTAAGAGSTP 411
Db 2281 ELSPVARQTSIQIGSSKAPBSGSDSTPSRPAQ-QPLSRPIQSPGRNISTISGRNGISPP 2339

QY 412 GNVLDVLTQYRGLSNVSVDENIGYTTTLANILRENPISLKG--SVGTGK--ETSRAFL 468
Db 2340 NK-----LSQLEPRTSSPSTASTKSSGSGKMSYSTSPGRQM 2373

QY 469 NOAVAGRAEDLPVQARLVLDLMDKNDTRIHFQED-WKITFLFIGNDLDFCNDLVHYSP 527
Db 2374 SQ-----ONLTKQ-----TGLSKNASSIPRSASAKGLMNQNGG---ANKVIELSR 2418

QY 528 QNFTDNGKALDILHAEPRAFAVNLVTLVLEIYNRELVOEKVYCPRMILRLSLCPVLKP 567
Db 2419 MSSTKSSGSESD--RSERP-----VLVRQSTRIKE-----APSPFLR-----2453

QY 588 DNSTELATLIEPNKKFQEKTHOLJSSGRYDTRREDTIVVQPFPEVNDPKTSEGLPDNS 647
Db 2454 -----RKLEBSASFESLSPSSRPASPTPSQAQTIVLSPSLPD-- 2490

QY 648 FFAPDGFHFSKSHDAASALNMNM---LEPV-----GOKTRHKEENKINICPNQVOP 699
Db 2491 -----NSLSTHSSVQAGCRKRLPPNLSPTIEYDNGRPAKHDLA-RSHSSPSRLP 2540

QY 700 FLRTYKNSMQHGHTWLPCEDRAPSAHPHSVHALRPADIQVVAALGDSLTAAGNGISKPD 759
Db 2541 INRS-----GTWKRHSKHSSSLPRVSTWRTRTSSSSISLASESSEBK-----AKSE 2587

QY 760 DLPDVTYQIRGLSYSGDGSLENTVTLPIILR-----EPN-RNLGYAVGTADADTNA 813
Db 2588 DEGRVNS-----ISGTQSKENOVSAKGTWRKIKENERSPTNSTSTQTVSSGATN----- 2636

QY 814 FLNQAIVGAKAEDLMSQVOTLMQMKD 840
Db 2637 -----GAESKTLIYQMAPAVSKTED 2656

RESULT 15
US-08-452-655B-2
; Sequence 2, Application US/08452655B
; Patent No. 578366
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; APPLICANT: ANAND, RAKESH
; APPLICANT: CARLSON, MARY
; APPLICANT: GRODEN, JOANNA
; APPLICANT: HEDGE, PHILIP J.
; APPLICANT: JOSLYN, GEOFF
; APPLICANT: KINZLER, KENNETH
; APPLICANT: MARKHAM, ALEXANDER F.
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: THILVERIS, ANDREW
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Witcoff, Ltd.
 STREET: 1001 G Street, NW
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20001-4598
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/452,655B
 FILING DATE: 25-MAY-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/289,548
 FILING DATE: 12-AUG-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/741,940
 FILING DATE: 08-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Kagan, Sarah A.
 REGISTRATION NUMBER: 32,141
 REFERENCE/DOCKET NUMBER: 1107,49964
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-508-9100
 TELEFAX: 202-508-9299
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2843 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-452-655B-2

Query Match 1.6%; Score 120.5; DB 1; Length 2843;
 Best Local Similarity 19.8%; Pred. No. 0.46;
 Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

34 GQIMPELTKNSPPCPNPKGVNMPKSKVSLKPSDIXFVAIGNLEIPDPGTDLEK-92
 1921 GQKPILOKOSTPP-----QSSKDI-----PDRGAATDEKL 1951
 93 QDWTERRQOYCM---GVMTLSDIIR-----YFSPSVPM-----123
 1952 QNPAIENTPVCFSHNSLSLSDIDENNKNENPIKETEPDPSQGEPSKQASGYAPKS 2011
 124 -----PVC---HTGKRVIPHDGAEDLWIOAELV-----RNM 152
 2012 FHVEDTPVCFSSRNSSLSLSIDSEDL---LQECISSAMPKKKPSRLKGDNEKHSFRNM 2068
 153 ---KENLOIDFO-----FDWKL---NVFFSNASQ---CYLCPSAQ 185
 2069 GGILGEDLTLDLKDIOIPDSEHGLSPDENFWDKAIQEGANSIVSLHQAAAAACLSRQ- 2127
 186 NGIAGGVDELMLGVLYLOEVEPRAFVNLVDSLSEVAEVSQYHGTWLSPPAPCECNCSEET 245
 2128 ---ASSDSDBILSLKSGISLGGSP---PHLTPDOEKPEFTSNK-----GPRILKPEKS 2174
 246 TRLAKVVMQWSY-----QAMNSLLASSRYSEQESFTVVFOFPFYEETTPSLHSEDPRLQD 300
 2175 TLETKIESSEKIKGKRYKSLITCKVNSNEISGMQKP-----LOANMPSISR 2226
 301 STTLAHLNMRMEPAGEKDEPLSVKXGRPMKCPQOESPYLFSYRNSNYLTRLOKPODKL 360
 2227 GRIMI-HIPG--VRNSSSTSPVS-KKGPLKTPASKSP--SEGQATTSPRGAKPSVKS 2280
 361 EV-----REGAEIRCPDKPS-----DTVPTSVRLKPADINVIGALGDSLTAGAGASTP 411
 2281 ELSPVARQTSQIGSSSAPRSRSGSRDSTSPRAPO-QLSNRPIQSPGRNNSISPRNGISPP 2339
 412 GNVLDVLTQYRGSLWSVSGDENIGVTTLANTLREFNPSLKGF-SVOTGK--ETSPNAFL 468

2340 NK-----LSQLPRTSSPSASTKSSGSGCMASYTSPGRQM 2373
 469 NQAVAGRAEDLPVQARLVDMKNDTRHIFQED-WKITLFLGNDLDFCNDLVHYSP 527
 2374 SQ-----QNLTKQ-----TGLSKNASSIPRSBSAKGLQNNNGN---ANKVELSR 2418
 528 ONFTDNIGKALDILHAEVPRAPNLVTVEIVLRELVOEKVYCPRMILRSLCPVLKPF 587
 2419 MSTKSSGSESD--RBERP-----VLVQOSTIKK-----ASPTLR-----2453
 588 DDNSTELATLIEFNKKFOEKTQILIESGRYDREDFTVVQPFENVDMPKTSEGLPDNS 647
 2454 -----RKLSEASAFESLSPSSRPASPTRSQAQTFVLSPLPD--2490
 648 FFAPDCHFSSKSHSRAALAMNM---LEPV-----GQKTHKFNKKNITCPRNOVOP 699
 2491 -----MSLSTHSSVQAGMKLPPNLSPTIEYNDGRPAKHDA-RSHSESPRL-P 2540
 700 FLRTYKNSMQGHGTWLPICRBRAPSAALHPTSVAHLPADIQVVAALGDSLTAAGIGSKPD 759
 2541 INRS-----CTWIKREHSKSHSSLPVSTWRTKRTGSSSILSASSESEK-----AKSE 2587
 760 DLPDVTTOYRGLSYSAGDGSLENVTTLPNILR-----EFN-RNLGYAVGTGDANDTNA 813
 2588 DEKHVNS-----ISGTKQSKENQVSAGTWRKIKENEPSPTNSTSQTVSSGATN-----2636
 814 FLNQAIVGAKAEDLMSQVOTLMQKMD 840
 2637 -----GAEKTLIYQMAPAVSKTED 2656

Search completed: January 6, 2004, 19:37:43
 Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2004, 18:14:14 ; Search time 100 Seconds
(without alignments)
3762.407 Million cell updates/sec

Title: US-10-054-691-2
Perfect score: 7766
Sequence: 1 MGIRPGIFLLELLILGQCT.....RCRRGRREDPPMSLRVAL 1458

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rv1rus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5624.5	72.4	1463	11	070320
2	5279.5	68.0	1450	11	054728
3	2270	29.2	423	4	081UP7
4	1392	17.9	270	4	096DP9
5	1145	14.7	394	11	09D4Y6
6	636	8.2	377	5	09N2Z4
7	635	8.2	186	11	08K255
8	628.5	8.1	382	5	023119
9	625.5	8.1	374	5	001300
10	616	7.9	398	5	08MX08
11	584	7.5	424	5	09VW80
12	515	6.6	447	5	09VW88
13	469.5	6.0	460	5	095QJ1
14	406.5	5.2	348	5	021799
15	360.5	4.6	528	5	P90862
16	360.5	4.6	565	5	081476

17	359	4.6	484	5	Q21798	Q21798 caenorhabd
18	314	4.0	471	5	Q9N4P5	Q9N4P5 caenorhabd
19	215	2.8	433	5	Q965M0	Q965M0 caenorhabd
20	159.5	2.1	941	16	Q97CZ3	Q97CZ3 clostridium
21	146.5	1.9	1755	16	Q9PC33	Q9PC33 xyella fab
22	145	1.9	2153	12	Q9Y0R5	Q9Y0R5 tula virus
23	141.5	1.8	799	4	Q8N265	Q8N265 homo sapien
24	141.5	1.8	873	5	Q01922	Q01922 caenorhabd
25	139	1.8	939	16	Q8CP76	Q8CP76 staphylococ
26	137	1.8	4306	11	Q9J179	Q9J179 rattus norv
27	136	1.8	685	16	Q8XST4	Q8XST4 ralsctonia s
28	135	1.7	1399	5	Q9V4R0	Q9V4R0 drosophila
29	134.5	1.7	16274	5	Q8IR22	Q8IR22 drosophila
30	133	1.7	1708	17	Q26769	Q26769 mechanobact
31	132	1.7	809	3	Q9HEA5	Q9HEA5 neurospora
32	132	1.7	1181	16	Q9PLB6	Q9PLB6 chlamydia m
33	131.5	1.7	1003	10	Q94856	Q94856 nicotiana t
34	131	1.7	465	12	Q90729	Q90729 human papil
35	130.5	1.7	891	10	Q9FZ40	Q9FZ40 arabidopsis
36	130.5	1.7	1165	5	Q01923	Q01923 caenorhabd
37	130.5	1.7	1545	5	Q9XX31	Q9XX31 caenorhabd
38	130.5	1.7	2792	5	Q814R2	Q814R2 plasmodium
39	130.5	1.7	2922	5	Q81PJ3	Q81PJ3 drosophila
40	130.5	1.7	3270	4	Q8WWW3	Q8WWW3 homo sapien
41	130	1.7	923	17	Q8U434	Q8U434 pyrococcus
42	129.5	1.7	1245	12	Q9Y709	Q9Y709 atelina her
43	128.5	1.7	974	10	Q8S905	Q8S905 arabidopsi
44	128.5	1.7	1003	10	Q9LP05	Q9LP05 arabidopsi
45	128.5	1.7	2765	11	Q9RIK2	Q9RIK2 rattus norv

ALIGNMENTS

RESULT 1	ID	070320	PRELIMINARY;	PRT; 1463 AA.
AC	070320;			
DT	01-AUG-1998 (TEMBLrel. 07, Created)			
DT	01-AUG-1998 (TEMBLrel. 07, Last sequence update)			
DT	01-MAR-2003 (TEMBLrel. 23, Last annotation update)			
DE	Phospholipase B.			
GN	PLB.			
OS	Cavia porcellus (Guinea pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.			
OX	NCBI_TaxID=10141;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Ctrl.(BPA)BR; TISSUE=Intestine;			
RX	MEDLINE=98256256; PubMed=9593672;			
RA	Delagebaudet C., Gassama-Diagne A., Nauze M., Ragab A., Li R.Y.,			
RA	Capdevielle J., Ferrara P., Fauvel J., Chap H.;			
RT	"Ecotopic Epithelial Expression of Guinea Pig Intestinal Phospholipase			
RT	B. Possible role in sperm maturation and activation by limited			
RT	proteolytic digestion";			
RL	J. Biol. Chem. 273:13407-13414(1998).			
DR	EMBL; AF045454; AAC40129.1; -.			
DR	InterPro; IPR001087; Lipase_GDSL.			
DR	InterPro; IPR005829; Sug transporter.			
DR	Pfam; PF00657; Lipase_GDSL; 3.			
DR	PROSITE; PS01098; LIPASE_GDSL_SER; 2.			
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.			
SQ	SEQUENCE 1463 AA; 162175 MW; 101C4969815F24B1 CRC64;			
Query Match	72.4%;	Score 5624.5;	DB 11;	Length 1463;
Best Local Similarity	73.5%;	Pred. No. 0;		
Matches 1074;	Conservative 150;	Mismatches 226;	Indels 11;	Gaps 7;
QY	1 MGIRPGIFLLELLILGQCTPOIHTSPRKSTLEGQIMPETLXNSPPCNPNKLGVMPSK 60			
DB	1 MGIRPGIFLLELLILGQCTPOIHTSPRKSTLEGQIMPETLXNSPPCNPNKLGVMPSK 59			

```

QY 61 SVHSLSKSDIKFVAALIGNELEPPDGTGDELEKODMTER-FOQVGMVTVLSDIIRYFSP 119
DB 60 SVHTLFPADIKLALAIQDMETPPDSGAVNLDTSRTKEKPMRGCMGMVTLSDIISHFN 119
QY 120 SVMPEVCHTGRVIPHDAEDLMIQAOELVNMKENIQLDPQDWKLIINFPNSAQCYL 179
DB 120 SVLLPTCPMRSAAVARGVEELFQAEELVSLKKNQLDFQODWKLIINFPNSAQCYL 179
QY 180 CPASQONGLAGGVDLGMVLDYQOEVPRAFVNLVDLSVABVSROYHGTWLS-PAPEP 238
DB 180 CPASHEGPMLSNMNDKAGILYHDEVPRAFNVLVDLFEVAMPRAHOGTMLSRSPEA 239
QY 239 CNGSEETRLAKVVMQWVSYOEAWSNLASSRSEOSFTVVPQPFYEETPLSHSEDPRL 298
DB 240 CGSGGSEKLDIVVMQWVSYOETWDSLASSSFNDQESFAVVPQPFYEVSPPV--EERPS 297
QY 299 QDSTTLAMHLNMRMEPAGEKDEPLSVHGRPMKCSQESPYLFSYNSNYLTRLQKPD 358
DB 298 QDPTTLALSLMNMKPKVQCKDEPFSTIERRPMKCPQESPYLFTYNSNYOURLKQR 357
QY 359 KLEVRGAEIRCPDQSDPVPPTSVMRLKPADINVGALGDSLTAAGAGSTGNTVDL 418
DB 358 QHREBGETEIRCPDQSDPPTSVMRLKPADIKVIGALGDSLTAAGAGSRPNTLIDL 417
QY 419 TOYRGLSMVSGDENIGTVTTLNLIREFNPSLKGFSVGTGKETSRYAFINQAVAGGRAE 478
DB 418 TEYRGLSMVSGDENIGTVTTLNLIREFNPSLKGFSVGTGKETSRYAFINQAVAGRAE 477
QY 479 DLEVOARLVLMKNDTRHIFQEDWKIITLFIGGNDLCDFCNDLVHYSPOFTDNIGKAL 538
DB 478 DLIPQARTVLDMKNHTSINFEBDWKIITVFIGGNDLCDFCSDPVTNSPENFTDNIRQAL 537
QY 539 DILHAEPRAFNVLVYELVNLRELQOEKRYCPRIILSLCEPVLKPDNDTELATLI 538
DB 538 DILHAEPRAFNVLVYELVNLRELQOEKRYCPRIILSLCEPVLKPDNDTELATLI 537
QY 599 EPNKKFOEKTHOLIESGRYDREDFTVVQPFENVMKPSGSLPNSFFAPDCPFSS 658
DB 598 DINKKOERHOLIESGRYDREDFTVVQPFENVMKPSGSLPNSFFAPDCPFSS 657
QY 659 KSHSRASALMNMLEBVQKTRHKEFNKINITCPNOVQPFRTYKNSMOGHTWLPOR 718
DB 658 KSHSRASALMNMLEBVQKTRHKEFNKINITCPNOVQPFRTYKNSMOGHTWLPOR 717
QY 719 DRAPSLHPTSVVALRPADIOVVAALGDSLTAAGGSKRDDLPTVTOYRGLSYSGGD 778
DB 718 DRAPSLHPTSVVALRPADIOVVAALGDSLTAAGGSKRDDLPTVTOYRGLSYSGGD 777
QY 779 GSIENVTTLNLIREFNRLTGYAVGTGANDTNAPLANQAVPAKAEMLSOVQTLMOQM 838
DB 778 GSIENVTTLNLIREFNRLTGYAVGTGANDTNAPLANQAVPAKAEMLSOVQTLMOQM 837
QY 839 KDDHRYNFHEDWKIVITVLIGGSDLCYCTDSNTYSANFVDHRLNALDVLHREVPRLVN 898
DB 838 KDDHRYNFHEDWKIVITVLIGGSDLCYCTDSNTYSANFVDHRLNALDVLHREVPRLVN 897
QY 899 LVDPLAPTIKROVFLGNPKDCPVQOASVLCNCTVTLRENOELARLEAFRATRSSMREL 958
DB 898 LVDPLAPTIKROVFLGNPKDCPVQOASVLCNCTVTLRENOELARLEAFRATRSSMREL 957
QY 959 VSGGRADTODSFVVLQPFQNTQLPYLADGLPDTSPFAPDCIHPNOKHPSOLARALMTN 1018
DB 958 VSGGRADTODSFVVLQPFQNTQLPYLADGLPDTSPFAPDCIHPNOKHPSOLARALMTN 1017
QY 1019 MEPLPSKETELDLRAEMPTTCPTONEPPLRTPRNSNYTYPIKPAIENMGSDELCTEMKA 1078
DB 1018 MEPLPSKETELDLRAEMPTTCPTONEPPLRTPRNSNYTYPIKPAIENMGSDELCTEMKA 1077
QY 1079 SNSVPTSVHOLRPADIKVVAALGDSLTTAVGAPRPNSSDLPTSMRGLSWSIGGDNLETH 1138
DB 1078 SNSVPTSVHOLRPADIKVVAALGDSLTTAVGAPRPNSSDLPTSMRGLSWSIGGDNLETH 1137
QY 1139 TTLPNILKKNPVLGPGSTSTWEGTAGLVAAAGARBDMPAQAMDVEMKNSPDINLE 1198

```

```

DB 1138 TTLPNILKKNPVLGPGSTSTWEGTAGLVAAAGARBDMPAQAMDVEMKNSPDINLE 1197
QY 1199 KDKMVLTLPLGVNDLCHYCNBPRAHATERYVGHITQOALDILSEELPRAPVNYE-VMELA 1257
DB 1198 MDKMLTLPLFGSNDLCHYCNBPRAHATERYVGHITQOALDILSEELPRAPVNYE-VMELA 1257
QY 1258 SLVQGGGKGC-AMLAQNNCTCLRHSSQSSLEKOLKKNMNLQHGSSFSYMHQYTOR 1316
DB 1258 GLHOGGGRCTALLPQOSTCSCLRHFPSSPVYQELKKTWNLSQSDMSRLSYOKTYOR 1317
QY 1317 FAVVQPFQNTLPLPNERGDTLTFPSSDCPFPSDRGAEMALAMNMNLEPVGKRTTS 1376
DB 1318 FAVVQPFQNTLPLPNERGDTLTFPSSDCPFPSDRGAEMALAMNMNLEPVGKRTTS 1377
QY 1377 NNETHSRALKKCPSPSPPLYTLKNSRLPDDAEEAPVLYANVPVPAAG---VGLVGI 1432
DB 1378 NNETHSRALKKCPSPSPPLYTLKNSRLPDDAEEAPVLYANVPVPAAG---VGLVGI 1437
QY 1433 IGVVWRCRGRGRREDPPMSL 1453
DB 1438 LAVSLMSFRROKKSPPSSV 1458

```

RESULT 2

```

ID 054728 PRELIMINARY; PRT; 1450 AA.
AC 054728,
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Phospholipase B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SD, TISSUE=Small intestine;
RX MEDLINE=98113187; PubMed=9442065;
RA Takeuchi H., Zolotarev F.N., Ting L., Urban T., Komatsu H.,
RA Hatanaka T., Okamoto M., Tojo H.;
RT "Identification of functional domains of rat intestinal phospholipase
RT B/lipase. Its cDNA cloning, expression, and tissue distribution.";
RL J. Biol. Chem. 273:2222-2231(1998).
DR EMBL; D63648; BA423813.1; .
DR InterPro; IPR01087; Lipase_GDSL.
DR Pfam; PF00657; Lipase_GDSL; 3.
DR PROSITE; PS01098; LIPASE_GDSL_SER; 2.
SQ SEQUENCE 1450 AA; 16108 MM; 45589808PD91F45 CRC64;

```

Query Match

68.0%; Score 5279.5; DB 11; Length 1450;

Best local similarity 70.0%; Pred. No. 0;

Matches 1012; Conservative 161; Mismatches 263; Indels 9; Gaps 6;

```

QY 5 PGVFL--LELLILGGCTPOHTSPKSLBEQOLMPELTKNSPPPCNPKLGVNMSKSV 62
DB 5 PGVFLGGLLLLLGGSPQIHSSGENTSQOQVFRITKNSSFPKPKLESLVLSKV 64
QY 63 HSLKPSDIKFAVAILGNELEPPDGTG--DLEK-QDWTERRPOQVCMGMVTLSDIIRYFSP 119
DB 63 HSLKPSDIKFAVAILGNELEPPDGTG--DLEK-QDWTERRPOQVCMGMVTLSDIIRYFSP 119
QY 120 SVMPEVCHTGRVIPHDAEDLMIQAOELVNMKENIQLDPQDWKLIINFPNSAQCYL 178
DB 120 SVMPEVCHTGRVIPHDAEDLMIQAOELVNMKENIQLDPQDWKLIINFPNSAQCYL 178
QY 1258 GLHOGGGRCTALLPQOSTCSCLRHFPSSPVYQELKKTWNLSQSDMSRLSYOKTYOR 1317
DB 1258 GLHOGGGRCTALLPQOSTCSCLRHFPSSPVYQELKKTWNLSQSDMSRLSYOKTYOR 1317
QY 1317 FAVVQPFQNTLPLPNERGDTLTFPSSDCPFPSDRGAEMALAMNMNLEPVGKRTTS 1376
DB 1317 FAVVQPFQNTLPLPNERGDTLTFPSSDCPFPSDRGAEMALAMNMNLEPVGKRTTS 1377
QY 1377 NNETHSRALKKCPSPSPPLYTLKNSRLPDDAEEAPVLYANVPVPAAG---VGLVGI 1432
DB 1378 NNETHSRALKKCPSPSPPLYTLKNSRLPDDAEEAPVLYANVPVPAAG---VGLVGI 1437
QY 239 CNGSEETRLAKVVMQWVSYOEAWSNLASSRSEOSFTVVPQPFYEETPLSHSEDPRL 298

```

Db 245 CKSEBITKSKAVMOWSVOEAMEDLLASSKFNKHETFAVVQSFSEV--ELPLERBSP 302
 QY 299 QDSTTLAMLMNMBEPAGEDEPLSYKGRPMKCPSESPYLSYRNSNYTLRLKOPD 358
 Db 303 QDSTTLALRLWNSMMEBVGKDGTLNEAKRTMKCPSESPYFTIRNSNYQAROLPEIG 362
 QY 359 KLEVEGAEIRCPDKPSDTPVTSVHRLKPADINVIGALGDSLTAAGASTPGNVLDVL 418
 Db 363 KFMKGKTKFTCPDKOPSDSIPTTVHRLRPADIKVIGAMGDSLTAAGASSTPGNVLDVL 422
 QY 419 TQYRGLSMSVSGDENICTVTTLANILREPNLSKGSVGTGKTSNPANLNOAVAGRAE 478
 Db 423 TQYRGLSMSVSGDEITETVTTLANILREPNLSKGSVGTGKTSNPANLNOAVAGRAE 482
 QY 479 DLVQARBLVDMKNDPRHIFROEDMKITLPIGNDLCCDFCNLDLVHNSPBNFNDNGAL 538
 Db 483 GLAAQAKVLSMKDKTTFQEDMKITLPIGNDLCCGNNLARSPTFTDNTKTAL 542
 QY 539 DILHAEPRAFVNLVTVLEIVNLRELYOEKVCPRMILSLCPVCLKPNDSTELATLI 598
 Db 543 DILHAEPRAFVNMVSTIEITPLRELPNEPKVCPMILSLCPVCLKPNDSTELATLI 602
 QY 599 EPNKKTQOEKTHOILIESGRYDTRDFYVVOPEPBNVDMPTSGLPDPSFPADCFHSS 658
 Db 603 ERNRQYOEETGLIESGRYDTRDDFTVVLQPMFENVMPRTLGLPDSFFADCFHFN 662
 QY 659 KSHSRASALMNMNLEBVGKTRHKEFNKINTCPNOVOPRLRYKNSNGHGTMLPCR 718
 Db 663 KTHARALALMNMNLEBVGKTRHKEFNKINTCPNOVOPRLRYKNSNGHGTMLPCR 722
 QY 719 DRAPSLHPTSVHRLPADIQVVAALGDSLTAAGISGSKPDDPDVTTQYRGLSYAGD 778
 Db 723 EKAPSAISPTSVTLRPADIQVVAALGDSLTAAGISGSKPDDPDVTTQYRGLSYAGD 782
 QY 779 GSLENTVTLNLIREFENRLTGYAVGTGANDTNAPLNOAVGAKADELMSQVOTLMQX 838
 Db 783 KPELENTVTLNLIREFENRLTGYAVGTGANDTNAPLNOAVGAKADELMSQVOTLMQX 842
 QY 839 KDHHRNHEHEDMKVITVLIGSSDLCOYCTDSNLKSAANFYDHLRNALDVLHREPVLVN 898
 Db 843 KNDTRVNFQDMKVITVMI GASDLCOYCTDSNLKSAANFYDHLRNALDVLHREPVLVN 902
 QY 899 LVDFLNPITIRQVFLGNPKCPVQOASVLCNVLTLRENSOELALEAFARSSREL 958
 Db 903 LVDFLNPITIRQVFLGNPKCPVQOASVLCNVLTLRENSOELALEAFARSSREL 962
 QY 959 VSGGRYDTRDFSVLQPFQNTQLPVADGLPDTSEFADCIHPNOKFHSOLARLMTN 1018
 Db 963 VSGGRYDTRDFSVLQPFQNTQLPVADGLPDTSEFADCIHPNOKFHSOLARLMTN 1022
 QY 1019 MLEPLGSKETTLDRAMPITTCPTQNEPRLTRPNSNYTPIKPALENMGSDPLCTEMKA 1078
 Db 1023 MLEPLGSKETTLDRAMPITTCPTQNEPRLTRPNSNYTPIKPALENMGSDPLCTEMKA 1082
 QY 1079 SNEVPTSVHQLRPADIKVVAALGDSLTAAGARPNNSDLPSTMRGLSMSIGDGNLETH 1138
 Db 1083 SNEVPTSVHQLRPADIKVVAALGDSLTAAGARPNNSDLPSTMRGLSMSIGDGNLETH 1142
 QY 1139 TTLPNILKKNPVLGFTSTWEGTAGLVAAEGARADMPAQAAMDIVERMKS PDINLE 1198
 Db 1143 TTLPNILKKNPVLGFTSTWEGTAGLVAAEGARADMPAQAAMDIVERMKS PDINLE 1202
 QY 1199 KDMKLVTLFGVNDLCHYCENPEAHATEYVQHIQOALDILSEELPRAFNVEVVELAS 1258
 Db 1203 KDMKLVTLFGVNDLCHYCENPEAHATEYVQHIQOALDILSEELPRAFNVEVVELAS 1262
 QY 1259 LVQGGGKCAM-LAAQNNCTCLRHSSSLEKOLKKNVNLQHGISFSFYWHQYOTREDP 1317
 Db 1263 LVQGGGKCAM-LAAQNNCTCLRHSSSLEKOLKKNVNLQHGISFSFYWHQYOTREDP 1322
 QY 1318 AVVVOPEFQNTLPLNERGDTDLTFPSEDCFHSRDRGHAEMALAMNMNLEPVGRKTTSN 1377
 Db 1323 AVVVOPEFQNTLPLNERGDTDLTFPSEDCFHSRDRGHAEMALAMNMNLEPVGRKTTSN 1382

QY 1378 NPTSPRAKLCPSPEPPLYTLRNSRLPDQAEAEVLYMAVPAAGVGLVIGITVY 1437
 Db 1383 NPTSPRAKLCPSPEPPLYTLRNSRLPDQAEAEVLYMAVPAAGVGLVIGITVY 1442
 QY 1438 WRCRR 1442
 Db 1443 WRTVR 1447
 RESULT 3
 Q81UP7
 ID 081UP7 PRELIMINARY; PRT; 423 AA.
 AC 081UP7;
 DT 01-MAR-2003 (TEMBLrel. 23, Created)
 DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
 DE Similar to phospholipase B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL, BC042674; AA042674.1; -
 SQ SEQUENCE 423 AA; 47354 MW; A20A193491BA780C CRC64;
 Query Match 29.2%; Score 2270; DB 4; Length 423;
 Best local Similarity 99.8%; Pred. No. 17e-160;
 Matches 422; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1036 MPITCPTQNEPRLTRPNSNYTPIKPALENMGSDPLCTEMKASNSVPTSVHQLRPADIK 1095
 Db 1 MPITCPTQNEPRLTRPNSNYTPIKPALENMGSDPLCTEMKASNSVPTSVHQLRPADIK 60
 QY 1096 VVAALGDSLTTAVGARNNSDLPSTMRGLSMSIGDGNLETHTTLPNLIKKNPVLG 1155
 Db 61 VVAALGDSLTTAVGARNNSDLPSTMRGLSMSIGDGNLETHTTLPNLIKKNPVLG 120
 QY 1156 STSTWEGTAGLVAAEGARADMPAQAAMDIVERMKS PDINLEKDKLVTLFGVNDLCH 1215
 Db 121 STSTWEGTAGLVAAEGARADMPAQAAMDIVERMKS PDINLEKDKLVTLFGVNDLCH 180
 QY 1216 YCENPEAHATEYVQHIQOALDILSEELPRAFNVEVVELASLVQGGGKCAMLAQNN 1275
 Db 181 YCENPEAHATEYVQHIQOALDILSEELPRAFNVEVVELASLVQGGGKCAMLAQNN 240
 QY 1276 CTCLRHSSSLEKOLKKNVNLQHGISFSFYWHQYOTREDPAVVQPFQNTLPLNER 1335
 Db 241 CTCLRHSSSLEKOLKKNVNLQHGISFSFYWHQYOTREDPAVVQPFQNTLPLNER 300
 QY 1336 GPDTLTFPSEDCFHSRDRGHAEMALAMNMNLEPVGRKTTSNFTSRKLCPSPEP 1395
 Db 301 GPDTLTFPSEDCFHSRDRGHAEMALAMNMNLEPVGRKTTSNFTSRKLCPSPEP 360
 QY 1396 LVTLRNSRLPDQAEAEVLYMAVPAAGVGLVIGITVWRCRRGRREDPMSLRT 1455
 Db 361 LVTLRNSRLPDQAEAEVLYMAVPAAGVGLVIGITVWRCRRGRREDPMSLRT 420
 QY 1456 VAL 1458
 Db 421 VAL 423
 RESULT 4
 Q96DP9
 ID 096DP9 PRELIMINARY; PRT; 270 AA.
 AC 096DP9;
 DT 01-DEC-2001 (TEMBLrel. 19, Created)
 DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)

DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE Hypothetical protein FLJ30866.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Brain;
 RA Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,
 RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
 RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
 RA Kikuchi H., Kanda K., Matsuo K., Nakamura Y., Sekine M.,
 RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
 RA Sugano S., Nagahara K., Masuno Y., Nagai K., Isogai T.,
 RT "NEDO human cDNA sequencing project."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK055428; BAB70920.1;
 DR Interpro: IPR001087; Lipase_GDSL.
 DR Pfam: PF00657; Lipase_GDSL; 1.
 DR Hypothetical protein.
 SK SEQUENCE 270 AA; 30869 MW; AC53CBCT5FD09523 CRC64;
 Query Match 17.9%; Score 1392; DB 4; Length 270;
 Best Local Similarity 100.0%; Pred. No. 2, 2e-95;
 Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 491 MKQDRIHFQEDMKITLFLFGNDLCPDNDLVHSPQNTDNGKALDILHAEVPAFV 550
 DB 1 MKQDRIHFQEDMKITLFLFGNDLCPDNDLVHSPQNTDNGKALDILHAEVPAFV 60
 QY 551 NLTVLEIVNLRELQYKRYKCPMLIRSCPCVLKRDNDSTELATLIEFNKKRQETHQ 610
 DB 61 NLTVLEIVNLRELQYKRYKCPMLIRSCPCVLKRDNDSTELATLIEFNKKRQETHQ 120
 QY 611 LISSGRYDREDFTVWVQPFENVDMKTSGLPDNSFPAPDCFHSSKSHRAAALMN 670
 DB 121 LISSGRYDREDFTVWVQPFENVDMKTSGLPDNSFPAPDCFHSSKSHRAAALMN 180
 QY 671 NMLEPVQKTRHKEFNKINITCPNOVQPLRTYKNSMOGHGWLPCRDRAPSALHPTSV 730
 DB 181 NMLEPVQKTRHKEFNKINITCPNOVQPLRTYKNSMOGHGWLPCRDRAPSALHPTSV 240
 QY 731 HALRPADIOVVAALGDSL 749
 DB 241 HALRPADIOVVAALGDSL 259
 RESULT 5
 Q9D4Y6 PRELIMINARY; PRT; 394 AA.
 AC Q9D4Y6;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DE 4930539A06Rik protein.
 GN 4930539A06Rik.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; Tissue=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Matsuzaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Okazaki H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schirml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Maehio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,
 RA Lyons P., Marchionni L., Maehima J., Mazzarelli J., Monbaerts P.,
 RA Nordone P., Ring B., Ringwald C., Rodriguez I., Sakamoto N.,
 RA Sasak H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitlaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohenki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK015993; BAB30072.1;
 DR MGI:19922406; 4930539A06Rik.
 SK SEQUENCE 394 AA; 44290 MW; E6D3ECL7E6F0888 CRC64;
 Query Match 14.7%; Score 1145; DB 11; Length 394;
 Best Local Similarity 59.8%; Pred. No. 1, 1e-76;
 Matches 234; Conservative 46; Mismatches 97; Indels 14; Gaps 7;
 QY 1 MGRLPGL--FLLELLLLGQGTQITSPKSTLEQ---LMPETLKNSPFCNPKLG 54
 DB 1 MELYPGVSPVGLLLLLGQSPQIHSSGENTLAWQSOQVFW--TLKRPFCCKPKLE 58
 QY 55 VNNPSKSVSLKPSDIFVVAIGNLEIPDPGTG--DLEK-QDTERPOOVCGWTVLTS 111
 DB 59 LSVLSESVSLKPSDIFVVAIGNLEIPDPGTG--DLEK-QDTERPOOVCGWTVLTS 118
 QY 112 DIIRYSPSPVMPVCHTKGRVIPHDAEDLMIQAEVLVNMKNENLQDFQDKLINVFF 171
 DB 119 DIIRHSPSVLMTPTGSPKGTAVHTTAEDLMIQAEVLVRLKONPQDPEKDKLITVEF 178
 QY 172 SNASQCLCPSAQNLGAAGVDELGVLDYLOEIPRAFNVLVDLSEVAESRQHGTM 231
 DB 179 SNTSQCHLCPSAQNLGAAGVDELGVLDYLOEIPRAFNVLVDLSEVAESRQHGTM 238
 QY 232 LSPAPPCNCSBETRLAVVMQWSYQEAAMNSLLASRSSEBSFTVFPQFFETPSL 291
 DB 239 FSPAPPCNCSBETRLAVVMQWSYQEAAMNSLLASRSSEBSFTVFPQFFETPSL 297
 QY 292 HSEDPLQDSTTLAMWLNRMMEPAGEKDEPLSVKGRPMKCSQSPYLFYSRNSNYLT 351
 DB 298 KRSSP--QDPTTLALINSMWMEPVQKGLNLTARKTMKCSSESPYLFYKNSNYQA 355
 QY 352 RLQKPDKLEVRGAEIRCPDKDPSPTVTS 382
 DB 356 RLKPYTKLQMKESSEFTCPDKNPSNSIPTT 386
 RESULT 6
 Q9NZ24 PRELIMINARY; PRT; 377 AA.
 AC Q9NZ24;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DE 4930539A06Rik protein.
 GN 4930539A06Rik.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Br1501 N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium."
 RL Science 282:2012-2018(1998).
 RN [2]
 SK SEQUENCE FROM N.A.

```

RC STRAIN=Bristol N2;
RA Maupin R., Bauer C.;
RT "The sequence of C. elegans cosmid Y65B4BR.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL/AC024847; AAF60857.1; -.
DR WormRep; Y65B4BR.1; CE25539.
DR InterPro; IPR001087; Lipase_GDSL.
DR Pfam; PF00657; Lipase_GDSL; 1.
KW Hypothetical protein.
SQ SEQUENCE 377 AA; 41400 MW; 2A93D3842C80E6A5 CRC64;

Query Match
Best Local Similarity 8.2%; Score 636; DB 5; Length 377;
Matches 132; Conservative 71; Mismatches 122; Indels 16; Gaps 8;

QY 376 SDTPVPSVRLKPADINVGALGDSLTAGAGSTGPNVDVLTQYRGLSMSVGDENIG 435
DB 37 SKNTPTSASVSRPTDIKVLGALDLSLAANGAGAPKGDPLAVILQYRGLAFQCGDHSID 96
QY 436 TTTTLNIIREFNPISLKGFSVGTGKETS-PNAF---LNOAVAGRAEDLPVQARRLVDM 491
DB 97 EHTVAVNLKKEFSPLMNGYSTGIG--SANYEVSKLNOVPGAEDIDITQARALVQII 153
QY 492 KNDTRIHFQEDWKITLFIGNDLDFCNDLVH--YSPQNTNIGKALDILAEVPRAF 549
DB 154 OSHKEIDYKTDKMLINVFICANDMCAYCNDRENGPSKAIWKQNVITAIQIDNLPTI 213
QY 550 VNLVTVLEIVNRELVOEKVYCPRMILKSLCPVLKFDNSTELATLIEFNKKFOEKT 609
DB 214 VSMTGMFDAMMLQI-DHDKYFCDGLHVE-CPCEKKNKDPNPDISAACHL---YMDAQ 268
QY 610 QLTESGRYDREDFTVVVQPFENVDMPKTS-EGLPNSFPAEDCPFFSKSHSRASAL 668
DB 269 EIDDSGIFPSTDTFTVQVDFNGITVPLKPDSEVNLDFAPDCFFSKGHANVAKIL 328
QY 669 MNMLEPVGOKTRHKEFN-KINITCPNOVQPFRLTYKNSM 708
DB 329 MNNIQVPGSKNQNVLSDPTILNCPDTCPCFFRTTKNSV 369

RESULT 7
Q8K255 PRELIMINARY; PRT; 186 AA.
AC O8K255;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Hypothetical protein.
DE Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Straube R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL/BC033606; AAH33606.1; -.
KW Hypothetical protein.
SQ SEQUENCE 186 AA; 21471 MW; 56153A626A22315 CRC64;

Query Match
Best Local Similarity 8.2%; Score 635; DB 11; Length 186;
Matches 116; Conservative 18; Mismatches 26; Indels 0; Gaps 0;

QY 1289 QELKVVNMLQHGISFSFYHQYTOREDFAVVVQPFQNTLPLNKGDTDLTFPSDCP 1348
DB 2 QELKVVNMLQHGISFSFYHQYTOREDFAVVVQPFQNTLPLNKGDTDLTFPSDCP 61

```

```

QY 1349 HFSDRGHAEALAMNNMLPVGKRTTSNNFTSHRAKLKCPSPESPYLTNRSRLPQ 1408
DB 62 YFSDRGHAEALAMNNMLPVGKRTTSNNFTSHRAKLKCPSPESPYLTNRSRLPQ 121
QY 1409 AEEPAEVLVAVPVAAVGLVVGIIIGTVVRCRGRRED 1448
DB 122 AEEPSNALYVAVPVAAVGLVVGIIIGTVVRCRGRRED 161

RESULT 8
Q23119 PRELIMINARY; PRT; 382 AA.
AC Q23119;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE W02B12.1 protein.
DE W02B12.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCB1_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Swinburne J., Alnough R.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=9069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 282:2018-2018(1998).
DR EMBL; Z66521; CAA91393.2; -.
DR WormRep; W02B12.1; CE03761.
DR InterPro; IPR001087; Lipase_GDSL.
SQ SEQUENCE 382 AA; 41772 MW; 6430145B8525CA3 CRC64;

Query Match
Best Local Similarity 8.1%; Score 628.5; DB 5; Length 382;
Matches 139; Conservative 57; Mismatches 119; Indels 37; Gaps 8;

QY 376 SDTPVPSVRLKPADINVGALGDSLTAGAGSTGPNVDVLTQYRGLSMSVGDENIG 435
DB 35 SKNTPTSASVSRPTDIKVLGALDLSLAANGAGAPKGDPLAVILQYRGLAFQCGDHSID 94
QY 436 TTTTLNIIREFNPISLKGFSVGTGKETS--PNAFLNOAVAGRAEDLPVQARRLVDMKID 494
DB 95 EHTVAVNLKKEFSPLMNGYSTGIG--SANYEVSKLNOVPGAEDIDITQARALVQII 154
QY 495 TRIHQEDWKITLFIGNDLDFC----NDLVHSPQNTNIGKALDILAEVPRAF 549
DB 154 OSHKEIDYKTDKMLINVFICANDMCAYCNDRENGPSKAIWKQNVITAIQIDNLPTI 213
QY 550 VNLVTVLEIVNRELVOEKVYCPRMILKSLCPVLKFDNSTELATLIEFNKKFOEKT 609
DB 214 VSMTGMFDAMMLQI-DHDKYFCDGLHVE-CPCEKKNKDPNPDISAACHL---YMDAQ 268
QY 610 QLTESGRYDREDFTVVVQPFENVDMPKTS-EGLPNSFPAEDCPFFSKSHSRASAL 668
DB 269 EIDDSGIFPSTDTFTVQVDFNGITVPLKPDSEVNLDFAPDCFFSKGHANVAKIL 328
QY 669 MNMLEPVGOKTRHKEFN-KINITCPNOVQPFRLTYKNSM 708
DB 329 MNNIQVPGSKNQNVLSDPTILNCPDTCPCFFRTTKNSV 369

RESULT 9
Q01300 PRELIMINARY; PRT; 374 AA.
AC Q01300;
DT 01-JUL-1997 (Tremblrel. 04, Created)

```


RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Maye A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mleishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong X., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Garxin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AEO0314; AAF9069.1; -
 DR EMBL: AY069174; AAL3919.1; -
 DR FLYBase: FBgn0036939; CG7365.
 DR InterPro: IPR001087; Lipase_GDSL.
 DR Pfam: PF00657; Lipase_GDSL.
 DR PROSITE: PS01098; LIPASE_GDSL.
 SQ SEQUENCE 424 AA; 48479 MW; 15DC6229188E171 CRC64;

Query Match 7.5%; Score 584; DB 5; Length 424;
 Best Local Similarity 38.5%; Pred. No. 8.2e-35;
 Matches 117; Conservative 57; Mismatches 116; Indels 46; Gaps 12;

QY 369 RCPDQPSDTVPTSVHRLKPADINVIGALGDSLTAAGAGSTPGANVLDVLTQYRGLSWSV 428
 DB 88 RSPER-----PTSVHRLKPADINVIGALGDSLTAAGAGSTPGANVLDVLTQYRGLSWSV 139
 QY 429 GDEENIGTVTLTANILREPRSLKSPVSGKSTSPAFNAQVAGRADLVQARRLV 488
 DB 140 GGQYDMRKVTLTNILKEFPNLYGATKQISTDRSDVLAELAMSKDPMKAVLV 199
 QY 489 DLAKNDTRIHFQDWMKLTIFIGANDLCPDGNLVHY-SPQNTD-----NIGKALDILHA 543
 DB 200 RRGQRDPVRVMTSDMKVTLFIIGNN---DCTDICYPEPEKVDHMERMLKTYRYLRD 256
 QY 544 EYPRAVNLTVLEIVNLRFL-----YQKKVYCPRMILRSICPVLFKDNSTSE 593
 DB 257 NVPRMLNITVPA---DNLRFLTNLTGILPICYG-----TLRECPCLM--GKQKQ 302

QY 594 LATLIEFNKKFOEKTROLIESGRYDREDFVVOVQFFENVDMPKTSEGLDPSFPAPDC 653
 DB 303 LDVLEGIMKRWIAKQDEIANRBEFNT-ETFTIVQSPQODPPRRRSQGTDRPSEDC 361
 QY 654 FHSSSKSHSBAALNNMLPEVQKT--TRKFEKNKINTCNOVOPPLRTYKNS 707
 DB 362 FHLISGRHAAANNSIWNMLLELGHKSGFATHLE---TFRCSEMRPFILITRENS 414

RESULT 12

Q9VNM8 PRELIMINARY; PRT; 447 AA.

AC Q9VNM8
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CG11029 protein.

OS *Drosophila melanogaster* (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephyroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gccayne J.D.,

RA Amande P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,

RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Maye A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Mleishina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong X., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,

RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AEO03611; AAF52285.1; -
 DR FLYBase: FBgn0031735; CG11029.
 DR InterPro: IPR001087; Lipase_GDSL.
 DR Pfam: PF00657; Lipase_GDSL.
 DR PROSITE: PS01098; LIPASE_GDSL.
 SQ SEQUENCE 447 AA; 51261 MW; BBD3F53DCCB35DEB CRC64;

Query Match

6.6%; Score 515; DB 5; Length 447;

Best Local Similarity 31.1%; Pred. No. 1.3e-29;
Matches 137; Conservative 73; Mismatches 155; Indels 76; Gaps 14;

QY 293 SEDRLQDS-----TTLAHMLNNMMEAGKDEPLSYKHGRPM 331
D 46 SYDRVRENGIQOYTIDIDQLRLFLNTRQTLTKMALNN-----IEALSSRGREG 96
QY 332 KCPQSESPYLFSYNSNVTRELQKPODKLEVEGAEIRCPDKP-SDTVPTSVRLKPAD 390
D 97 K-----LQAP-----VSKVPLCLCPNNNRSFSPSTISHLRPGD 131
QY 391 INVIGALGDSLTAAGAGSTPGNVLVLTQYRGLSWSVGSDENIGVTTLANILREPNPS 450
D 132 IDITAAAGDSLISANGILIS--NNAI DMINEERGLTFSGGGLANRRFVTLPTILKIFNPK 189
QY 451 LKGSVGTGKETS-PNAPLQAVAGRAEDLPVQARLVLMKRDTHIQEDMKITLF 509
D 190 LYGFVAVNSLIVNRRSLNIAEPMISRDLPQARVILDLRRDRVDMKRWKLLTVY 249
QY 510 IGGNDLDFGNDLVHY-SPQNFDT---NIGKALDILHAEPRAVNLVTVLEIVNREL 564
D 250 VGNNDI---CSDLCMTPOSFLDQHARDLQARFLRDHVPRLILVVPNIPLVLT 306
QY 565 YQEKRYVCPMILASLCPVLFKPDNSTELATLIEFKKQKQKTHQIESGRYDREDPT 624
D 307 MTKVPLQC-FVYHVGCHCLINDRLNTEFNERMDTLTRMQOLMEIARLPEF-RQDFA 364
QY 625 VVVOFFENVDMPTSGSLPDNSFPAPDCFFFSKSHSRASALMNNMLEVQCK---T 680
D 365 IVAPMLTKTLAPLPDGSIDMREFSHDCFFHSORGHAIISNLMNSMLLPDDQKPPSV 424
QY 681 TRHKFNKINITCNPQVQFL 701
D 425 VPBLFE---RVVCPTAEQPYI 442

RESULT 13

Q950L1 PRELIMINARY; PRT; 460 AA.

AC 0950L1 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, last annotation update)
DE Hypothetical 51.9 kDa protein.
GN P13H8.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;
OC Rhabditiida; Pelodertinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RA Ding H.;
RT "The sequence of C. elegans cosmid P13H8.";
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RA Waterston R.;
RT "Direct Submision.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U23133; AL042483.1; -
DR WormBep; P13H8.11; CE29264.
DR InterPro; IPR001087; Lipase_GDSL.
DR Pfam; PF00657; Lipase_GDSL.1.

KM Hypothetical protein.
SQ SEQUENCE 460 AA; 51861 MW; 5D4108B1657319E1 CRC64;

Query Match 6.0%; Score 469.5; DB 5; Length 460;
Best Local Similarity 29.6%; Pred. No. 3.2e-26;
Matches 120; Conservative 72; Mismatches 174; Indels 39; Gaps 14;

QY 358 DKLEVEG-AEIRCPDKPEDTVPTSVHRLKPADINVIGALGDSLTAAGAGSTPGNVL 416
D 43 DSNIEYPMGFNPNCP--QPIKSSSEVHQHPQIOGVALGDSVSAQAAS--SILD 98
QY 417 VLTQYGLSWSVGSDENIGVTTLANILREFNSLGFVGTGKETS-PNAPLQAVAGR 476
D 99 LFPQFQVSVFTGDDVYLNQATFINIPRFAPRIKG---GSSDFQKFDYFMALPGSF 155
QY 477 AEDLPQARLVLMKRDTHIQEDMKITLIEGGNDLDFGNDLVHYSPQNFDTNIGK 536
D 156 SSELPRQALKLASTLKKLGRETSDTWKFVNIFIGHNLCNNNETTGPPEFGSLHS 215
QY 537 ALDILHAEPRAVNLVTVLEIVNRELVOEKKV--YCPMILRSICPVLFKPDNSTEL 594
D 216 ALITIGTNPKVFN--IMPIVNIHQAHLSKFC-EFSHRKTCSCIFELNEKEYQ- 270
QY 595 ATLIEFKKQKQKTHQIE--SGRYDREDTVVVOFFENVDMPTSGSLPDNSFPAP 652
D 271 ---NIGKQDEQNEVEQFNQKGNSSFFAVVIAPADLKSIP-LKKNQPNIGLLAD 325
QY 653 CFFFSKSHSRASALMNNMLEVQCKTRHKFNKINIT-----TCNPQVQFLRYKNS 707
D 326 CFHLSPLAHIDIAKQIKWIGLFEPIDQKT---ITNLSVGFDFVCPVPECPLRTIONS 361
QY 708 MGHGTMPLPCRD---RAPSLHPTSVHRLPADIOVVALGDSL 748
D 382 ENCE---PSREKRFILVPSAFASSPGLSMPTMIVFFLFGIAL 422

RESULT 14

Q21799 PRELIMINARY; PRT; 348 AA.

AC Q21799 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, last sequence update)
DT 01-NOV-1996 (TEMBLrel. 01, last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, last annotation update)
DE R07B7.8 protein.
GN R07B7.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;
OC Rhabditiida; Pelodertinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B.R.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z75955; CAB00118.1; -
DR WormBep; R07B7.8; CE06270.
DR InterPro; IPR001525; C5_DNA_meth.
DR Pfam; PF00657; Lipase_GDSL.
DR POSITE; PS00095; C5_MTASE_2; 1.
SQ SEQUENCE 348 AA; 39489 MW; 91F52DC5139363E0 CRC64;

Query Match 5.2%; Score 406.5; DB 5; Length 348;
Best Local Similarity 32.2%; Pred. No. 1e-21;
Matches 117; Conservative 66; Mismatches 133; Indels 47; Gaps 11;
QY 1060 IKPAIENWGSDFCTEMKASNSV-----PTSHQARPADIKVVALGDSLTAAGARPN 1113

Db 15 VKNALP-----KYLNSIDPHEIYDPDVMMIKPHIRIYIGAMGDSLT--IGSRAB 61
QY 1114 NSDDLPTSMWGLSWSIGGDNLEHTTLPLNLIK-----KENPYLLGPSTWEG-TAGLNV 1168
Db 62 NIVGQORIPGNAFFTGMDVEVDRLVYINIFRIIAEKTGNKLFGGSTGIDYBENTLVN 121
QY 1169 AAGCARADMPAQAAMDIVERMKNSPDINLEKMWLVTLFIGVND---LCHYCENPEAHLA 1225
Db 122 AIGGMSDOLIRAKELVSRITANKENIENDMKLWSLWGTNDVGLGRLEDPPIR-V 179
QY 1226 TEYVOHIQALDILSEELPRAFYNNVEVWELASLYOGQGGKCMALAAQNNCTCLRHSQS 1285
Db 180 DEYKSHIEKGLYLKENLPRTIVISVGMFPAQLQEAQ-----SILKCKRA 226
QY 1286 LEKQELKXVNMNQHGSFSFYHQYTOR---EDFANVYQPF-FQNTLTPLNREGDTLT 1341
Db 227 RIVENQCKLD-DLSDGGRNVSYDFQNNHFNNSNDFYVQPFATEYTDSDYRDEHGXNPT 285
QY 1342 FFESEDFHFSDRGHAEMALALMNNMLEPVGRKTTSNFTSRAKLCSPSPSYLYTLRN 1401
Db 286 FYASDLFHLKSKFGHAYLAKHYMLNLPVGVGKTKRADLGDTKXIYELNEKNCILITVGN 345
QY 1402 SRL 1404
Db 346 SKM 348

RESULT 15

P90862 PRELIMINARY; PRT; 528 AA.
ID P90862;
AC P90862;
DT 01-MAY-1997 (T-EMBLrel. 03, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Hypothetical protein F36A2.9a.
GN F36A2.9a.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Leonard N.?
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology";
RL Science 282:2012-2018(1998).
DR EMBL; Z81077; CAB03072.2; -.
KW Hypothetical protein.
SQ SEQUENCE 528 AA; 59421 MW; 1B34202465E16261 CRC64;

Query Match 4.6%; Score 360.5; DB 5; Length 528;

Best Local Similarity 29.9%; Pred. No. 5.5e-18;
Matches 103; Conservative 59; Mismatches 134; Indels 49; Gaps 13;

QY 370 CPDKDSPDTVPTVHRLKPADINVIGALGDSLTAAGNAGSTPGNVLDVLTQYGLSMVSG 429
Db 130 CPRIKTELTLGNVGNISPEIDITIIAMGDALTGIGLWPN---ADI--EFGASFPFG 183
QY 430 GDEINIGTITLANILREFNPSLKGFSVGTGKETSPPNAFLNQAAYAGRAEDLPVOARRLYD 489
Db 184 GDSITIDGLTIPILREBFSPKLVGVSHGAGADL-PNHQLNVAVTGATTEDELPGQARRLTR 242
QY 490 LMKNDRIHQEDWKIITLFIGNDLCPDNLVHVSPPQNTNIGKALDILHAEPVPAF 549
Db 243 RLKLELDELPHNEMIMIIITIGTEELCSRCGPSY-----DNIRKAIHLOIEIPKA- 294
QY 550 VNLVTLVLEIYNLELVOEKVYCPRMILRLCPVLKFDNSTELATLIEFNKKPQE--- 606

Db 295 --LVVLLGPVHVSFHEQKS-----NLKARCAC-----SRDQTE-GFMYDVSRKSWKVMR 342
QY 607 KTHQLLESGRYDTRDPTVYVQPFENVDMPK-TSGLPDNSSFADPCHFFSKSHSRAA 665
Db 343 DVQKFEVNG-----VTARETFGMISYPMVLTITSRYPSGLFIRDKPLNRRGHNYAT 393
QY 666 SALMNM-----LEPVGQKTRBKFEKNKINITCPNOVOPFLRTYKN 706
Db 394 KMLNRLIGDLYNLSSATL-----SQDNYFCPSVGCPIFRITYEN 433

Search completed: January 6, 2004, 18:59:17
Job time : 108 secs

THIS PAGE BLANK (USPTO)